The role of polymorphism and life history trade-offs during range expansions

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy by

Vincent Arthur Keenan

September, 2018

Abstract

Understanding how fast a species invades new habitat or expands its range is of critical importance. As a result of climate change, the number of species expanding their native range is increasing. More exotic species are being inadvertently released into habitats unprepared for their arrival. Evidence suggests that expansions and invasions are occurring at faster rates due to the evolution of dispersal. However, few mathematical models exist that consider how the rate of spread is influenced by dispersal life history.

This thesis addresses the gap in understanding using mathematical models to make predictions about invasive spread under a dispersal-reproduction trade-off. Using deterministic models and simulations, we build upon existing theories of anomalous invasion speeds, extending the cases for which they are known to exist. We have also examined the literature and performed a meta-analysis to determine the prevalence of dispersal trade-offs in nature that can be used to validate existing theory.

We show that anomalous invasion speeds are a robust phenomenon that are possible within populations that express a general degree of polymorphism, but at most 2 phenotypes determine the speed of invasion. This result was found to persist when interactions among species were non-neutral, moreover, those interactions do not influence the invasion speed. Invasion speeds are shown to decrease in the presence of another species, even when one of the 2 phenotypes that determined the invasion speed is out-competed to the point of negative net growth. Anomalous invasion speeds are also shown to exist when a species reproduces sexually. When that species is diploid, we find the invasion speed is sensitive to how the heterozygote expresses its dispersal and reproductive traits.

Overall, we found that accounting for the life histories related to dispersal has surprising and robust effects that alter the dynamics of biological invasions.

In heartfelt memory of Paula Agnes Elliott

1990 – 2018

She shines brighter than ever; we shall never be dancing

in the dark.

Declaration

Contributions for the contents are as follows

Chapter 2 contains work that has been prepared for publication in a peer reviewed journal. Vincent A. Keenan (VAK) and Stephen J. Cornell (SJC) jointly designed the model. VAK conducted the simulations and analysis that resulted in the determination of the invasion speed. SJC developed the arguments for the trade-off curve determining the invasions speed and conducted the analysis contained in Appendix A. VAK and SJC both contributed to drafts of the manuscript for which this chapter was based on. SJC gave guidance and edited the final draft of the manuscript.

The contributions of chapter 3 were as follows. VAK and SJC both jointly designed the model. VAK conducted the simulations and analysis that resulted in the determination of the invasion speed. VAK wrote the content of the chapter. SJC gave guidance with analysis and simulations and helped edit drafts of the chapter.

The contributions of chapter 4 are as follows. VAK and SJC developed the initial ideas for the chapter. VAK, SJC and Raj Whitlock (RW) developed the search and inclusion criteria used to gather the data. VAK and Amy Eacock (AE) collected the data used within the analyis, with AE aiding in the clarity of the inclusion criteria. VAK conducted the analysis with aid from SJC and RW. VAK, SJC, and RW interpreted the data. VAK wrote the content of the chapter.

SJC helped edit the chapter.

The contributions of chapter 5 are as follows. VAK and SJC both jointly developed the model. VAK conducted the simulations used within the chapter. VAK wrote the content of the chapter. SJC helped edit the chapter.

Both the introduction and conclusion were written by VAK and SJC helped edit them.

Acknowledgements

I would like to thank my secondary supervisors Jane Hill and Ilik Saccheri for helpful discussion about aspects of my work, particularly for the soft advice for finishing my PhD. I would like to thank the Ecology, Evolution, & Behaviour department at the University of Liverpool for being welcoming and providing an environment that was pleasant to work under. I would like to thank all the people within the PhD offices and post doc offices for all of the times they were able to distract me from doing what I was supposed to – all work and no play. I'd like to thank ACCE DTP for giving me ample opportunities that expanded my expectations of a PhD and beyond. Many thanks are also deserved of the non-academic staff within the Institute of Integrative Biology, they have worked diligently behind the scenes so that we can all continue doing what we enjoy.

I would like to thank my family for always being supportive and giving me my room back whenever I went home, despite knowing they are enjoying life without my constant carry on. I would like to thank the lads of Liverpool Hibernia Football Club for giving me something to do on a Saturday afternoon, evening, and often wasting a Sunday morning in bed...

I would like to thank some people personally. Jamie Alison for saving me from a mould and misery factory masquerading as a flat when I first started, if he didn't alleviate me from there I may have packed the whole thing in at an early stage. I would like to thank Becky Jones, Georgia Drew, Chris Corbin, and Lingzi Wang for being great housemates. Those times spent together will never

leave me, along with the jelly that I still occasionally find behind my ears. I would like to thank Jack Pilgrim, he introduced me to my football team, and has always been there when I needed it most – he is my ballast. Thanks to Gabbi Pedra, we were two analytical PhDs in a sea of empiricists keeping each other from isolation. I would like to thank Adam Fisher for the many times we either in the pool, or out on our bikes – particularly when I was writing up – which helped me relax when I needed it most. I would also like to thank Raj Whitlock for his advice and patience when helping me conduct my meta-analysis, his effort was invaluable to that chapter.

Very special thanks to Amy Eacock for helping my complete my meta-analysis, I simply couldn't have done it without her. I would also like to thank her for constantly being my sounding board when I was at my darkest points.

Finally, I would like to thank my supervisor Stephen Cornell. I would not be the scientist I am today because of his solid guidance, remarkable patience, and judicious advice. No one has had a greater impact on my way of thinking about things beyond my parents than he has, and for that I am eternally grateful. He has helped me to not only grow as a scientist, but as a person. Here's to future collaborations!

Contents

List of Figures

List of Tables

Chapter 1

Introduction to range expansion and invasions

1.1 Introduction

Throughout this thesis we will be exploring range expansions and biological invasions. It is therefore important to make a concrete definition of what each of these terms actually imply, particularly since the terminology is debated – especially within invasion literature [185, 10, 191]. Mathematically, the process of both range expansions and invasions are similar, however, there are distinct nuances in their definitions. There is no precise definition of either range expansion or invasion, and much disagreement. Therefore, throughout this thesis, we shall use the following definitions

Biological invasion – The process by which a species expands from its native range into favourable habitat with the only impediment being the limits imposed by its own traits.

Range expansion – The process by which a species expands its native range into new habitat made favourable by changing climatic conditions.

Both of these definitions share similarities, but the main difference is how favourable the habitat is to the species. During a range expansion, species will be resident at the very edge of what the species finds physically sustainable and the rate of expansion is determined by the advance of favourable conditions. Whereas during a biological invasion, the species will often be introduced to favourable habitat with no limit and the rate of advance is limited by the traits of that species.

1.1.1 Why study range expansions & invasions?

Understanding the process of range expansion and invasion is one of the most important topics within ecology, evolution, and epidemiology. It is the general scientific consensus that changes in global surface temperatures are likely to exceed 1.5 °C by the end of the 21st century according to all predictions [131]. An observed effect of a warming climate on all species is that their native ranges are shifting as a consequence [21, 132]. As climate warming becomes more intense, species will be forced to adapt to new conditions, or perish, if they cannot keep pace with the advancing climatic envelope [177]. Those species which do keep pace will inevitably encounter other ecosystems for which they may, or may not, have catastrophic consequences, such as the sea urchin (*Centrostephanus rodgersii*) in South Eastern Australia which has had catastrophic effects for the reef ecosystem[110].

Economic concerns

Species which invade new territory or expand their range can have serious impacts on local economies. One key tenet within the invasion literature is the description that "invasion" implies a negative impact. When a non-native species invades into a new habitat it can have negative consequences for human interests. Invasive species have been attributed the responsibility of $97 billion in damages from the period of 1906-1991 in the United States of America [139, 140], and in Canada estimates of $13.3-34.5 billion Canadian dollars per year [31]. This is an acute problem within food production, with the population of the earth expected to grow to 9 billion people this mid-century [63]. Invasive species disrupt the food

production process, from the brown rat (*Rattus norvegicus*) – native to East Asia [112] – destroying grain [3], or from introduced crop pests lowering yields [40, 140], or from water hyacinth (*Eichhornia crossipes*) reducing the quality of fish within Lake Victoria [92].

Invasive species also change the landscape, altering the water cycle or increasing fire risk. In New Zealand native tussock grasslands are efficient at retaining moisture from precipitation, which play a crucial role in human water supplies, but are being replaced by non-native pasture grasslands [114, 82]. A similar invasion in South Africa with non-native species of the genera *Pinus, Eucalyptus,* and *Acacia* reduced water yields, but levels returned when they were removed, but at great expense [187]. The case in South Africa is especially pertinent, where the city of Cape town experienced major water shortages during the summer of 2018 [1]. Fire risk can also be increased when non-natives displace indigenous species while amassing more flammable biomass that can cause wildfires resulting massive economic damage [16, 37].

Health concerns

As species expand their range, they can be accompanied by pathogens harmful to humans, and livestock. As a result of climate change, mosquito borne diseases are being reported in higher altitudes and deeper within the Sahara as temperatures are cooler during foraging times [50, 149]. One study predicts warming temperatures provide suitable conditions for a range expansion of the tick species (*Ixodes scapularis*), which is a vector for Lyme disease [128]. More species are invading newly optimised habitat, accompanied by potentially malignant zoonotic diseases. There is an increase in potential for a vector population to reach a threshold where a host shift to humans is possible, such as Dengue fever in the American Tropics [199]. Any such disease outbreak would put intense pressure on health infrastructure through inoculation, hospital capacity and staffing, and economic loss from incapacitated workers.

Biodiversity concerns

One of the largest threats to biodiversity is invasive species. Researchers disagree on the level of blame associated to invasives for extinctions of natives, but the threat has been identified [27, 67]. When an invasive or range expanding species moves into habitat already occupied by another resident species, there is inevitably an effect on the ecosystem. The effects of the newcomer are numerous, from competing with the residents directly [162], disturbing nutrient cycles [11], disrupting food webs through predation [18], or the introduction of pathogens [38]. Red squirrels (*Sciurus vulgaris*) are currently declining in the United Kingdom as a result of the grey squirrel (*Sciurus carolinensis*) from a combination of competition and the squirrel parapoxvirus [152, 178]. The decline of the red squirrel in the United Kingdom is an example of when an invader can displace a species through pathogens.

1.2 Important traits of range expanding species

The effect of range expansions and biological invasions are a priority for researchers within ecology, evolution, and epidemiology. The topics discussed above are only a sample of the broader implications that the complex process of expansion has on an ecosystem. It is therefore critical that we understand what key factors influence this process. Important aspects that contribute to a successful range expansion or invasion can be catergorised as

1. dispersal

2. reproduction

3. variation within a species

4. species interactions

5. environmental interactions

which act together to influence how successful an expansion is, and the speed at which it is successful [78, 102, 22]. Although we have defined the list above, for the purpose of this thesis, we shall be focusing on items 1-4 and introducing what is currently known about these terms. Although we do not explicitly examine environmental interactions, we shall explain what is known about them in the context of these four points.

1.3 Invasions and dispersal

Dispersal is intrinsic to any species invading or expanding its range, and thus, it is critical that we understand this process. Currently, dispersal is the least understood aspect of range expansions due to its complexity [155, 163]. There is an established definition of dispersal which is

"Dispersal is the movement of an individual from site of birth to a site of reproduction (natal or pre-breeding dispersal) or its movement between successive site of reproduction (post-breeding or simply breeding dispersal) It is the main mechanism leading to gene flow within and between populations."[28]

It is this definition of dispersal which we shall use within the entirety of this thesis. The act of dispersing is considered to comprise of three stages: emigration, transfer, and immigration [150, 7]. Although all organisms disperse in some way, the motivations for dispersing are not universal, nor is there any strict pattern why any particular strategy emerges.

One element arises from the environmental conditions. Dispersal allows an organism to select favourable habitat by assessing available resources and the quality of such resources, from potential breeding sites, to the availability of nutrients [170]. This ability also allows for an organism to sense when conditions are deteriorating and vacate to more favourable locations [87, 13, 143]. Environmental conditions can also extend to other organisms within the habitat which have pressures on dispersal, such as predation and parasitism [167]. In the Rocky

Mountains in the United States of America, mayfly larvae (*Baetis bicaudatus*) adjusted their dispersal patterns to become nocturnal to avoid predators [117].

1.3.1 Genetics of dispersal

Variation among individuals influences dispersal, from behavioural to physiological aspects of the process. Many species develop different physiological traits to aid with dispersal, such as the development of wings, seed morphology, or stored nutrients for dispersing [15]. Such traits are largely determined by genetics, but there is growing evidence that plasticity plays a large role too [150, 155, 29]. There is evidence of a genetic role for dispersal in the expression of genes responsible for flight apparatus in the cricket (*Gryllus firmus*) [204, 147, 35]; and too within the butterfly Glanville fritillary (*Meletaea cinxia*) through the *Pgi* allele associated with flight capability [75, 127, 126]. However, the discovery of these genes that influence dispersal does not imply they are solely responsible for dispersal propensity. Results from studies examining the heritability of dispersal are mixed, much of the literature is unbalanced focusing on birds and insects which ranged from 0 to 1, with an average of 0.35 [155]. Heritability here is defined in the narrow sense, which is given by

$$h^2 = \frac{Var(A)}{Var(G) + Var(E) + 2Cov(G,E)},$$

where $Var(A)$ is additive genetic variance, $Var(G)$ is the variance of genotypic effects, $Var(E)$ is the variance of environmental effects, and $Cov(G,E)$ is the covariance between genotypic and environmental effects. The denominator is collectively known as the phenotypic variance.

A fundamental question about any genetic trait is whether its heritability is governed by a single gene, a few genes – oligogenic effects – or many genes – polygenic effects [155].

Since it is generally accepted that dispersal has some form of genetic basis, then it is those genes which are subject to evolution. Some of the best examples

of the evolution of dispersal are from range expanding or invading species. A well studied and famous example in Western Australia is the cane toad (*Bufo marinus*), which has been invading at an accelerating rate causing devastation to the local ecosystems when they arrive. Evolution has acted through a spatial means which has been given the term *"spatial sorting"*. Individuals at the invasion front have greater dispersal ability and mate with toads that have similar traits, which acts as a feedback loop increasing dispersal – the *"Olympic village effect"* [159]. Toads were recorded with longer legs, and had more directed movement patterns than toads deeper within the core [137, 135]. Both the cone-head bush cricket (*Conocephalus discolor*) and Roesel's bush cricket (*Metrioptera roeselii*) evolved to have longer wings within their macropterous forms – long winged phenotype – at their respective range edge [175].

So far we have discussed circumstances when dispersal has been favoured, but dispersal can also be selected against. If the landscape which a species finds itself in is a mosaic of favourable and unfavourable habitat – such as a classic meta-population – then dispersal has been found to be selected against [64, 118]. It has been theorised that Allee effects should also select against dispersal, since a dispersing population that suffers negative growth won't be able to sustain itself through local extinctions [181].

Local adaptation can suffer due to larger dispersal rates. If a species is dispersing then its adaptation to local conditions can be detrimental to survival [71, 58]. An effect known as *mutation surfing* is a consequence of greater dispersal. Genes which are maladapted to new local conditions are present within a dispersing species which mate with others at the range front, much like spatial sorting, increasing their frequency. If the landscape is a meta-population, then the consequences are slower expansion, or local extinctions [180, 51].

1.3.2 Dispersal measurements

Due to the complexity of the dispersal process, empirical measurements that are accurate and reliable are notoriously difficult to obtain, particularly during an

invasion or range expansion. What we can never know is the exact position of the invasion front or range edge, moreover, if any individuals found there are in transit. If the organism is in transit we cannot know precisely where it originated, and thus if it is travelling towards the range core, or the range edge. To undertake any survey of a species is impractical and is likely to be fraught with problems. To address this, proxies of dispersal are often used instead of actual distances to provide a measure of dispersal propensity. Proxies that are used can include wing length, mobility, thorax weight for flight muscles, and discrete winged phenotypes within insect species [165, 86, 75, 175]; stalk height, wing loading, terminal velocity, and seed size-weight ratio in plants [176, 57, 41, 189]; and tracking breeding locations in the form of philopatry and dispersers within many mammal species [65]. The number of different measures of dispersal are wide ranging, which makes comparison within, and among, species seriously difficult.

Since empirical measures are difficult, the hurdle of capturing realistic dispersal has been addressed using modelling techniques. The most classic case of modelling dispersal is by using Brownian motion, which assumes that a species makes small local movements throughout its lifetime. Classic results in the theory of invasions have been modelled using dispersal as a diffusion process, able to predict the speed of advance [129, 166]. However, the diffusion assumption has given mixed results when applied to real data. The current favoured approach is to use what is known as a *dispersal kernel* which is a probability distribution of the distance an individual will travel from its breeding location to its reproductive location. Under certain conditions, the classical diffusion case is a Gaussian dispersal kernel, which along with exponential distributions have been popular in the past [28]. Recently, so called *"fat-tailed"* dispersal kernels have been used to account for greater variation in long-distance dispersal events – which are not captured by Guassian or exponential distributions. Fat tailed distributions are not exponentially bounded, and thus can account for variation and mixing [106]. Predictions from dispersal kernels of this type have been found to match empirical data for wind-dispersed seeds very well [26]. Other methods that account for

dispersal in meta-population models is the use of a dispersal rule. To connect different patches, there is some probability of dispersal between patches generated from a distribution [76, 182].

It is clear that dispersal is a significant factor when investigating range expansions or invasions. Dispersal is difficult to quantify due to the complexity of the process, furthermore, when evolution is acting on a species there can be simultaneous effects from the environment and genetic architecture acting to favour, or punish dispersal. The evolution of a species is recorded biologically within the organism's life-history, which when accounted for, has strong implications for range expansions and invasions.

1.4 Life histories and trade-offs during invasion

The biological state of any organism is for the large part determined by the life history of that species, which during an invasion or range expansion is under intense selection pressure. Life history theory has primarily focused on what are known as $r - K$ traits, where r is associated with reproduction, and K with competitive traits. Theory has shown that r-selected traits are categorised as maximising reproductive potential in the form of early maturity, many small young, short lifespan and large reproductive effort; whereas K-selected traits are categorised as late maturity, few large young, and small reproductive effort [171]. The definitions of both r and K traits remain somewhat subjective, and difficult to quantify outside of theoretical studies. Within plant species, it has been noted that r - strategists are often the best invaders since many invasions happen in disturbed habitat void of competitor species [145].

However, during invasions and range expansions dispersal is a large component of the process and is considered a costly investment [12]. Therefore, it is more appropriate that dispersal should be included in the process as $r - K - D$ - selection. To date there is only one study which examines $r - K - D$ - strategies which found that investment in dispersal and reproduction were heaviest at the

range front at the expense of K - selected traits, whereas in the range core, it was $r - K$ selection strategies that dominated at the expense of dispersal [17].

As we have seen in section 1.3, dispersal is subject to selection pressures – if there is an increase in dispersal propensity, it has to come at some expense. There are two types of cost borne out: according to life history theory there is a physical trade-off between investment in dispersal apparatus and some other trait, or alternatively, as an indirect – such as behavioural – trade-off with some aspect of fitness for the organism [12].

Direct physical trade-offs

When a species incurs a physical trade-off between dispersal and another trait, it is in the form of energy diverted to other resources – such as locomotory dispersal apparatus, or energy divergence – that another trait must experience a reduction in development. The most commonly reported dispersal trade-off is with fecundity, but other traits are also known to incur a the costs of dispersal. The most documented cases are within cricket species (*Orthoptera*), which also notably experience dispersal polymorphism. Many studies have shown that trade-offs are expressed within crickets through two phenotypes, a flight capable macropterous form, and a flightless brachypterous form. The macropters have been shown to have larger flight capable wings, but at the expense of smaller ovaries; while the brachypters have small, or no wings, but have larger ovaries [147, 66, 123, 204, 202]. Dispersal polymorphism of the aforementioned kind is not only restricted to crickets, but has been reported in tobacco thrips [124] , chinch bugs [56], and aphids [205]. One study has found that fighting ability within leaf beetles (*Callosobruchus subinnotatus*), with smaller wings had larger mandible lengths which aid in confrontations with conspecifics over territory and mates [6]. In the speckled wood butterfly (*Pararge aegeria*), a trade-off between flight propensity and reproduction was found while expanding their range. It was found that thorax sizes (where flight muscles are located, a proxy for dispersal) in core populations were smaller, while having larger abdomens (where eggs are located, and a proxy for reproduction); whereas at the range edge the opposite

was true [86].

Although dispersal trade-offs are predicted to be found in a variety of species, some are not present, even within the same taxonomic group. Previously we mentioned that a dispersal trade-off was found within the speckled wood butterfly, while the Glanville fritillary, another butterfly, does not experience the same trade-off. Several studies into dispersal-reproduction trade-offs within Glanville fritillarys consistently lacked evidence of any such relationship [153, 154, 75].

Physical trade-offs with dispersal are also widely seen in plant species, with some displaying a strong dispersal polymorphism in different seed dispersal strategies. Plants have a large variety of adaptations for survival that many other organisms do not possess, such as dormancy, or clonal growth in the form of ramets – parts of the plant body which grow, or break off from the parent to become independent. The dispersal characteristics of a plant are very different for say, a butterfly, because they are sessile. Therefore, it is the progeny, or seeds, which actually disperse. Many plants possess dispersal polymorphism in the form of seed dispersal, for example, the desert annual (*Gymnarrhena micrantha*), the only species within the genus, produces wind dispersed seeds, and subterranean seeds. The wind dispersed achenes allow for long distance dispersal to colonise further territory, while the subterranean seeds populate the area nearby the parent plant [168]. The species *Gymnarrhena micrantha* is a member of the *Asteraceae* family, which has shown evidence of a dispersal trade-off between germination timing of wind, and non-wind dispersing achenes [43, 89, 138]. Seeds which germinate early have a distinct advantage to those that germinate later and can be seen as a competitive trait, and evidence of a dispersal-competition trade-off.

The dispersal strategies within plants are numerous, as are suspected trade-offs. Seeds which are dispersed through foraging by predators, known as *endozoochory*, must invest in seed coatings which can resist the corrosive effects of digestion at the expense of nutrients within that are required for development [174, 23, 52]. One hypothesis for dispersal-competition trade-off that is often

invoked is seed number vs. seed size. The concept of the dispersal-competition trade-off is that smaller, lighter seeds disperse further at the expense of developmental nutrients, whereas larger, heavier seeds do not disperse as far, but have more nutrients within which offer a competitive advantage [196, 184, 88]. Another trade-off within seed strategies is that of dormancy, which can be viewed as dispersal through time. One interpretation of dispersal through time is the effect of a dormant seed awaiting the conditions for emergence before germinating, as if it had only just arrived – since while it is dormant it is not using, or competing for, any resources. Dormancy is a common trait within desert plants, often becoming active with the arrival of rain [68]. The possibility of fire has a strong selective pressure on seed dormancy. When a habitat produces enough biomass, which implies strong competition for resources, there is a greater risk of fire during warm, dry conditions. Therefore, a successful strategy has been to invest in dormancy which responds to fire as a cue to germinate when competition is expected to be less [93]. The trade-off is between dormancy, or dispersing and germinating early to compete for resources. Naturally, an individual can begin reproduction earlier if it is not dormant and produce progeny for the next generation immediately, but a dormant individual can wait until conditions are better for a greater chance of successful reproductive effort.

Indirect trade-offs

When there is an indirect trade-off, dispersal can be accompanied by a cost to fitness. Within an organism it is often found that variation within a species is accounted for by discrete polymorphism, in particular a *dispersal polymorphism* whereby different phenotypes have large differences in traits. An example of dispersal polymorphism is when the decision to disperse is delayed, while other immediately disperse. Many cooperative breeding species display dispersal polymorphism in the form of delayed dispersal [97]. The benefits of dispersing allow the organism to escape competition for resources, but the cost can be reduced fitness from potential mortality during dispersal, or time spent searching for a mate, acquiring resources, and evading predators [201]. On the other hand, those that delay dispersal can benefit from diverting more resources into body condition for

later dispersal, or remaining in the relative safety of a natal patch, or access to mates if they rise in a hierarchy [98]. With or without dispersal polymorphism, the main costs of dispersal are paid by decreased fitness from direct mortality while dispersing, and from lost reproductive success by time spent dispersing, for example nomads.

We can see that the traits which dispersal can trade-off with are not simple, easily identifiable features. Much of the non-dispersal traits can be confounded by other aspects linked with fitness, which implies that empirically identifying a dispersal trade-offs with a high degree of certainty is tenuous. Identifying a trade-off in one species does not necessarily imply that the same trade-off can be found within the same species class, as we have seen with the peppered moth and Glanville fritillary. When attempting to uncover a dispersal trade-off, such as within the Glanville fritillary, it is not necessarily absent, but the trade-off is within some other aspect that we are unaware of, or potentially absorbed by many traits [12].

Theoretical research in trade-offs

The difficulties in studying dispersal trade-offs empirically means that any evidence is focused in a small collection of study organisms, or is tenuous at best. Therefore, much of our understanding of how dispersal trade-offs affect population dynamics has been progressed through theoretical studies. Much of current research addresses indirect trade-offs and focuses on the cost of dispersing from direct mortality [12]. The majority of models have focussed on evolutionary stable rates of emigration from patches within simulations and meta-population models [141, 85, 142]. Few studies have examined how the evolution of dispersal is incorporated, and less how a trade-off is incorporated into predictions [136, 8, 182, 183]. Non-physical trade-offs that involve decisions by individuals are much harder to quantify, and thus theoretical predictions about the costs associated with such behavioural decisions are difficult to test.

Theoretical research in the form of mathematical models have been devel-

oped to help explain the evolution of dispersal, and associated trade-offs. Early work by Hamilton and May [72] has shown that dispersal is advantageous to avoid inbreeding depression, but at the cost of some mortality during dispersal. More recently, new mathematical models have been developed to assess costs of dispersal. One form of dispersal trade-offs that has had increasing attention is that of a dispersal-reproduction trade-off. The probability that dispersal is an increasing function of the energy associated to traits that aid dispersal is high, particularly for locomotory apparatus [148]. Despite this, studies often assume a simplistic approach that some proportion of resources is allocated to a particular trait with no reasonable justification [94, 17]. Using the framework presented by Hamilton and May [72], Weigang & Kisdi conducted a study and found that the shape of any dispersal-reproduction trade-off curve was paramount to determining the evolutionary trajectory of any species [192]. Several studies have shown that the shape a trade-off curve can indeed have major consequences on evolutionary dynamics and should be incorporated into theoretical studies [151]. It has also been shown that if a dispersal trade-off does exist, then theoretically there is a trajectory that, depending on how intense selection is, will always end in dispersal polymorphism [192, 96]. Therefore, dispersal trade-offs give rise to evolutionary dynamics that can lead to dispersal polymorphism, which in the context of invasions and range expansions where intense and often rapid evolution occurs, could play a major role in the dynamics. We have arrived at significant point, whereby dispersal trade-offs and dispersal polymorphism could potentially be a key mechanism for successful range expansions and invasions – which are drastically under studied.

Studies that accept the evolution of dispersal and include dispersal trade-offs have shown that their inclusion in theoretical models has non-negligible effects on the population dynamics. Almost all studies which allowed for adaptive dispersal dynamics showed significant increases in the rate of expansion [137, 136, 55]. Few studies exist where trade-offs are used within range expansions, but that their inclusion has dramatic effects. Elliott and Cornell have shown that a simple dispersal reproduction trade-off in the form of dispersal polymorphism can increase

the speed of invasion if dispersal and reproductive abilities are sufficiently different [47, 48]. Theoretical models often predict the speed of invasions or range expansions to be less than is observed, but almost all studies assume rigid dispersal terms [78]. However, when allowing for the evolution of dispersal the trend indicates that predictions are faster than those of the classical models.

1.5 Species interactions

1.5.1 Inter-specific competition

Species interactions during range expansions or invasions are rarely considered in studies due to high levels of complexity already present during the process. If we exclude abiotic dispersal barriers, then competitive interactions from other species are a simple, and reasonable, explanation for limits to expansion [102]. It has been shown that range borders where a species is actively competing against another have stable range borders over ecological time [20, 101]. In Costa Rica, two genera of tropical song birds, *Catharus* and *Henocorhina*, aggressively responded to calls of the other at the point where their ranges met, resulting in a stable range for both species [90]. Competition between species is theoretically predicted to slow invasions if competition is weak enough, and stop them if strong enough, although investigations are rare [129, 108, 77]. There is evidence of slower invasions as a result of competition in real systems. The cane tibouchina (*Tibouchina herbaceae*), is an invasive species on the Big Island of Hawai'i that quickly occupies areas of open space when they become available [4]. It was found that when the cane attempted to invade forest communities, the established species reduced the rate of invasion. The same process has also been seen in California where native perennials slow the advance of invasive exotic annual grasses [158]. However, competition does not always slow expansion. In California resistance from a native ant community did not affect the spread of Argentinian ants (*Linepithema humile*).

1.5.2 Intra-specific competition

Competitive interactions within the same species are more difficult to assess than those from a competing species. Intra-specific competition is also known as density-dependence, and is often included within theoretical studies. Many models include density-dependence in the form of logistic growth for single species models [166], and Lotka-Volterra dynamics for multi-species models, which include a combination of intra-, and inter-specific competition [47, 129]. The theory of kin competition implies that in such instances evolution favours dispersal away from conspecifics to avoid competition with individuals that have the same alleles [83, 142, 72].

1.5.3 Predator-prey interactions

Interactions between predator and prey, or prey and pathogen/parasites can play a large role in range expansions. Predator-prey interactions during invasions and range expansions are often not studied, and thus predictions are not generally known, but a study by French & Travis has shown that they can influence spread [54]. Specialist predators or pathogens will naturally have ranges limited by the distribution of their respective hosts [102]. If a species range expands and overlaps with another predator, then this has the potential to slow or even stop an expansion by reducing population growth – and ultimately of the predator too [84]. Predators or pathogens are also capable of causing an expansion through selection for dispersal to escape predators, particularly when spatio-temporal dynamics become unstable which can select for dispersal [195, 143]. Predators and pathogens can also reduce local densities, which in turn increases intra-specific competition, which selects for dispersal [102].

1.5.4 Mutualism

Not all interactions are antagonistic – species that benefit mutually can influence range expansions. One of the greatest mysteries of expansion is known as *"Reid's paradox"* where Reid could not explain the rapid advance of English oak (*Quercus rober*) preceeding the Holocene Ice Age after observing the average dispersal

distance of seeds falling from their parent trees. The most widely accepted theory, postulated by Reid, of the rapid rate of expansion is the mutualism between fruits from the oaks providing sustenance for foraging species, and those uneaten fruits germinating further from the parent than would drop from the parent tree [144, 25, 24]. Various models incorporating long distance dispersal have attempted to account for this using fat-tailed dispersal kernels for rarer long-distance events [136, 25]. However, it must be noted that this is still an open problem, and the methods used in the solution provided by Clark and colleagues [25] have not been reliable in making general predictions for the spread of other invasions. This is in part due to the dispersal kernels requiring accurate parameterisation. Other examples of mutualisms aiding invasions are many species of plants that only began spreading after their pollinators arrived [146].

1.6 How range expansions and invasions are modelled

As we have noted throughout this chapter, range expansions or invasions have complex interacting factors that make empirical experiments difficult and impractical. Therefore, much of the research is conducted using theoretical methods in the form of simulation studies, or mathematical models. Using theoretical models is a more practical method of understanding range expansions, which can then be used to guide empirical research to test what factors influence expansions more directly. Models of range expansions and invasions have been subject to study since pioneering work by Fisher and Skellam [53, 166]. In their work, Fisher and Skellam both found simple mathematical expressions for the speed of invasion that were dependent on the dispersal and reproductive abilities of the species. Since their work, recent studies have found that the assumptions from their original models are too simplistic and do not capture the nuances of the other factors we have been discussing already. One particularly pertinent issue for any theoretical study is how dispersal is modelled. Different methods require particular assumptions, some of which are not valid in every system, resulting in large variation between methods [78]. Therefore, finding some form of generality

within predictions remains elusive.

1.6.1 Continuous time and space

The classical approach to investigate invasions has been the use of reaction-diffusion partial differential equations – PDEs – to model this process in continuous space and time. The previous work by Fisher and Skellam is an example of when this approach has been used in the past. This model has been used alongside empirical data to determine the invasion speeds of musk rats (*Ondatra zibethica*) in Europe, cereal leaf beetles (*Oulema melanopus*) in North America, and the small cabbage white butterfly (*Pieris rapae*) in North America [5]. However, the predicted rate of spread of leaf beetles was 20 times slower than the observed speed, indicating that other factors must be influencing the spread. The rate of invasion of grey squirrels in the United Kingdom has been modelled using reaction-diffusion equations in a Lotka-Volterra competition framework between competing grey and red squirrels and an invasion speed was predicted [129]. The predictions for the invasion speed were close, but it was found that the population dynamics did not reflect the displacement of the red squirrels, until it was discovered that disease dynamics also played a role [178]. Reaction-diffusion models also account for the population dynamics, as well as dispersal. These models can accommodate a variety of different population demographies associated with growth, such as an Allee effect. It has been shown that Allee effects dramatically slow the rate of spread [107]. The model used by Elliott and Cornell to model the invasions of a population with dispersal polymorphism, and more specifically, a dispersal-reproduction trade-off, is of the reaction-diffusion family [47].

1.6.2 Discrete time, continuous space

Reaction-diffusion models are powerful tools, but are not suitable for all types of expansions or invasions, particularly if the species has non-overlapping generations. A more natural model for organisms that have non-overlapping generations are integro-difference models, which are discrete in time, but continuous in space. Many univoltine insects fall into this category. The use of parameterised dispersal kernels has been used extensively in such models. These models are often

simplified to one dimension and the landscape is assumed to be homogeneous. In doing so, dispersal only relies on the distance between parent and offspring. Therefore, the probability of finding an individual at location x is the sum of the probabilities of all other locations y that an individual produced at y disperses to location x [78]. The difference between diffusion models and dispersal kernels is the generality of the probability distribution. It is in using dispersal kernels that rare long-distance effects can be included within models in the form of fat-tailed kernels. However, a serious problem with such kernels is that the rate of spread is sensitive to the behaviour at the tails of the distribution, where data is often lacking [78]. The real benefits of such models come from including age structure of any invading species. It has been shown that when age structure is taken into account, with periodic and stochastic effects, that invasions are predicted to be slower than when it is not [125]. Integro-difference equations, like reaction-diffusion, can also accommodate competitive interactions during invasions, which is more evidence that competition is predicted to slow invasions – a result that is robust to changes in modelling framework [77, 108].

1.6.3 Discrete time and space

When developing models, sometimes the available data or information about a species may not be continuous enough to use either framework above, in which case discrete time and space models can be used. If the data on an invasion is collected in discrete time periods for particular locations, such as annually, then a discrete time and space model would be appropriate. Using discrete instead of continuous space simplifies the model and makes it easier to analyse and simulate, especially if sexual reproduction is being investigated. When simulating, a continuous time and space model that accounts for sexual reproduction can become undefined when there is no mate present and a division by 0 occurs. The dynamics of discrete space and time models have been shown to be qualitatively similar to their continuous time and space counterparts [36]. Although useful for including sexual reproduction within models, there are few studies which examine discrete time and space models in the context of invasions, where integro-difference approaches are commonly preferred. Models of this kind have been used to study

predator-prey dynamics while incorporating Allee effects [122].

1.6.4 Meta-population and individual based models

Often the landscape for which a species is invading or expanding into is not homogeneously suitable, but rather fragmented into favourable or hostile patches. A natural way to address this issue is by using meta-population models which treat a species distribution not as a single population, but smaller connected populations. Meta-population models were first introduced to address habitat fragmentation by Richard Levins and championed by Hanski [105, 74]. Many species are expected to live within a meta-population due to anthropogenic land use, and changing climatic conditions [74]. Meta-population models are incredibly versatile, able to account for changes in sexual systems, include environmental and temporal heterogeneity, account for evolution through adaptive dynamics, and include interactions – all of which cover the majority of heterogeneity within dispersal and life histories [102, 175, 73]. However, the versatility of these models lack the simplicity that mathematical models can achieve. Each situation must be parameterised for the species and environment, so any conclusions are restricted to the scenario being investigated. These models are often simulated and therefore can only illustrate how scenarios are likely to end without informing us which parameters are important. Whereas through mathematical analysis, we can find which parameters are likely to govern the behaviour of the model, and under which values these parameters take. However, methods to study meta-population models have been developed [130]. Meta-population models have been used to model range expansions, whereas invasions are more commonly modelled using mathematical models. Meta-population models can be used to investigate the rate of expansion while accounting for other demographic features, and has been successful in many cases. It has been shown through meta-population models that the silver spotted skipper butterfly is expected to expand its range [175, 39].

Individual based models (IBMs) have been used to investigate rage expansion in a similar way to meta-population models. The difference between these approaches is that IBMs track individuals, whereas meta-populations consider

the populations. IBMs have the same benefits and downsides as meta-population models in that they are versatile, but lack the simplicity that mathematical models can deliver. IBMs are popular alternatives to meta-population and mathematical models because individual variation can be captured. These models have been used extensively to model range expansions [182, 45, 17, 179].

1.7 Overview of this thesis

We have introduced the motivations for studying range expansions and invasions along with key processes that determine their success and expansion rates. We have found that both the evolution of dispersal and dispersal trade-offs have had large impacts on the studies they have been included in, but rarely are they invoked in the study of invasions. Furthermore, no general theories exist that provide predictions for the effects that evolutionary consequences have on invasions. We shall aim to address the areas that are least understood, that is, how the evolution of life history under a trade-offs and dispersal polymorphism can affect range expansions and invasions, with a particular focus on the rate of advance.

1.7.1 Chapter 2

Previous research by Elliott & Cornell has shown that a dimorphic species that is under a dispersal-reproduction trade-off can give rise to faster invasions – known as *"anomalous invasion speeds"* [47]. However, in nature there are likely to be more than two phenotypes within a population. A natural question to ask is what happens when there are more than two? Will we still see anomalous invasion speeds? If so, what contributions do we expect from the other phenotypes? Are there any general predictions we can make from mathematical analysis? In this chapter we shall generalise this results from two phenotypes – later referred to as *strains* – to a general degree of polymorphism with N strains. We shall use a deterministic model that is homogeneous in time and space to derive the speed of invasion of the species, and conditions under which these faster speeds are realised. We shall show that all of the results from [47] hold for the N-strain

case, and that 2-strains is a special case. Furthermore, that in an N-strain system the invasion speed is determined by at most 2 strains within that population, while the remaining strains do not contribute to the speed. We shall also derive a geometrical argument that illustrates that the shape of the dispersal-reproduction trade-off curve determines whether faster invasion speeds are realised.

1.7.2 Chapter 3

Chapter 2 examined the case when a species invades vacant favourable habitat, but this is often not the case in nature. A species is more likely to encounter a competitor while invading, but what effect would this have on a polymorphic population? Would anomalous invasion speeds still occur? If so, does the resident affect the invasion speed like it does in other studies? Does polymorphism aid, or hinder invasion? In this chapter we seek to extend the previous results to the case when a competitor species is resident where the polymorphic species is invading into. We use the same deterministic model as in chapter 2, but modified to accommodate the resident and restrict polymorphism to the dimorphic case. We shall derive conditions for each species to invade the other, and show that anomalous invasion speeds are still attainable in the presence of a resident. We will also show that anomalous invasion speeds still persist even when one of the effects of the resident is to cause one of the strains to have negative growth.

1.7.3 Chapter 4

We stressed throughout the introduction that empirical evidence for dispersal trade-offs is lacking. Therefore, we would like to ask: does the information in the literature support theoretical predictions that dispersal trade-offs exist? If so, are there any predictors that indicate where we might find them? Do some traits have stronger trade-offs than others? To answer these questions, we shall conduct a systematic review of the literature for studies which have data that can be used to determine the existence of dispersal trade-offs and conduct a meta-analysis on these data to derive a conclusion. Through this, we will also gain an indication of how current research on trade-offs is conducted, such as organisms used, what conditions experiments are conducted under, the sex of the individuals used. We

will find that evidence for dispersal trade-offs does exists, furthermore, we shall find that the factor that consistently determines whether a trade-off is present is physical investment in dispersal apparatus. We will also show the contributions to the trade-off from fitness, sex, species class, climate conditions and environmental lineage.

1.7.4 Chapter 5

Finally, we will further investigate how robust anomalous invasion speeds are by changing the reproduction method from clonal, to sexual. Species that reproduce clonally and those that reproduce sexually have very different systems for passing genes to the next generation. A natural question is then, does the reproductive change stop the expression of anomalous invasion speeds? If they still persist, then what affect does genetic recombination play? We answer these questions with a discrete time and space model that examines a diploid species that experiences a dispersal-reproduction trade-off through gene expression of different alleles. Each homozygote will be either strong at dispersal, and weak at reproducing, and *vice versa*. We shall then investigate the effect that the heterozygote has when there is dominance for either trait, has the strengths of both homozygotes, the weakness of both homozygotes, and an intermediate of either trait. We will also establish that these solutions permit travelling wave solutions, and draw comparisons between our previous chapters.

Chapter 2

Anomalous invasions speeds of highly polymorphic populations

2.1 Introduction

We shall begin this thesis by investigating the invasion of empty habitat by a polymorphic species which experiences a dispersal reproductive trade-off. Classic work by Fisher and Skellam suggests that, after a transient period, invasions should proceed at a constant speed that is determined primarily by dispersal and population growth rate [53, 166]. While their model makes many simplifying assumptions — deterministic, logistic population growth with dispersal represented by Brownian motion — in theory their results should give a robust first approximation for invasion by a single monomorphic species – i.e. where all individuals have the same traits – with short-range dispersal into uniform habitat where local populations are not too small [200]. However, empirical invasions often do not conform to these theoretical predictions. For example, after the last glacial period tree species colonised Europe much more rapidly than predicted by short-ranged dispersal (the so-called "Reid's paradox", [166]), and cane toads (*Rhinella marina*) are advancing in Northern Australia at an accelerating rate, rather than at a steady speed [137, 144].

One possible reason for the departures from Fisher and Skellam's results is that real species contain multiple strains which can mutate and evolve. Dispersal polymorphism in particular has attracted much interest, as higher dispersal ability confers a clear advantage when colonising new habitat. Empirical data suggest that newly established populations contain individuals with elevated dispersal capabilities [164, 55], and the accelerating cane toad invasions are accompanied by an evolution of dispersal-related traits such as leg length [137]. The tendency of more dispersive strains to be found at the vanguard of an invasion has been dubbed "spatial sorting" [159].

However, invasions depend on population growth as well as dispersal. A higher investment in dispersal ability can imply a lower investment in traits related to reproduction [155], and a sufficiently strong trade-off could reverse spatial sorting — a highly dispersive strain cannot contribute to the invasion if it is infertile. Moreover, such a trade-off can itself give rise to very unexpected behaviour. Elliott and Cornell showed that, if a species exists in two strains (one a superior disperser, and one more fecund), then it can invade at a significantly faster, "anomalous", speed than would be predicted for either strain in isolation [47]. Simulation studies show that accelerating speeds can occur in the presence of a dispersal-fecundity trade-off [17], but there is currently no theory that predicts the biological conditions under which different invasion scenarios (e.g. spatial sorting, constant or accelerating speed, etc.) would be expected for polymorphic populations.

Here, we shall develop a theory for invasions by a species consisting of many strains, in which there is a trade-off so that strains with higher per-capita population growth rate have lower dispersal. We will show that the shape of the trade-off curve critically determines whether the invasion speed continues to accelerate or approaches an asymptote. The trade-off curve also determines whether the invasion speed equals that of one of the constituent strains, or whether the speed is anomalous, i.e. is faster than any single strain on its own. We will derive a geometric approach to illustrate the dependence of the speed of invasion on such

a trade-off. In all cases, we shall find that the asymptotic speed is determined by the traits of at most two of the constituent strains. We find that spatial sorting does not occur in all cases — for example, the most dispersive strain does not necessarily lead the invasion — and that the invasion speed is not necessarily determined by the most dispersive or the most fecund strain, or the one which, in isolation, would invade at the fastest speed. Our model is simple but generic, so our results should be a good qualitative description of the outcome of invasion in a wide range of biological systems.

2.2 Mathematical analysis

We shall extended the spatially explicit Lotka-Volterra partial differential equation model used in Elliott and Cornell [47] from 2 strains to the general case of N strains. The model is temporally and spatially continuous and homogeneous. We will only consider explicitly one spatial dimension, since the homogeneity of spatial conditions would produce an identical result in two or higher dimensions. We consider a haploid species that is polymorphic at one locus, where each genotype is expressed as a distinct phenotype. Each phenotype then differs in its dispersal, reproductive, and intraspecific competitive traits. We assume density independent birth and that mortality is affected by intraspecific density dependence. The population reproduces asexually with a constant proportion mutating at birth from their parental genotype into another genotype. Diffusion is modelled by Brownian motion. We henceforth refer to the distinct genotypes as "strains". The equations that govern the above dynamics are given by

$$\frac{\partial n_i}{\partial t} = \overbrace{D_i \frac{\partial^2 n_i}{\partial x^2}}^{Dispersal} + \overbrace{r_i n_i (1 - \sum_{j \in S} C_{ji} n_j)}^{Lotka-Volterra growth} + \overbrace{\sum_{j \in S \setminus \{i\}} \mu_{ji} n_j - \sum_{j \in S \setminus \{i\}} \mu_{ij} n_i}^{\substack{Mutation\ dynamics\ from\ population\ i \\ into\ populations\ j\ and\ vice\text{-}versa}}, \quad \forall i \in S, \tag{2.1}$$

where the elements of the set $S \equiv \{1, 2, ..., N\}$ label the strains within the population. The quantities n_i, r_i, and D_i denote respectively the population density, intrinsic per-capita population growth rate, and diffusion constant for strain $i \in S$. The intraspecific competition term C_{ji} describes the competitive effect of strain j on strain i. A fraction $\frac{\mu_{ij}}{r_i}$ of all offspring of strain i mutate at birth into

strains j. All parameters are assumed to be positive in order to be biologically relevant or, for the case of C_{ji}, so that intraspecific interactions are competitive rather than cooperative.

We shall consider a population in which all strains initially coexist in the semi-infinite landscape $x < 0$ and subsequently invade unoccupied habitat $x > 0$. If each strain is present in the landscape in isolation – which would require $\mu_{ij} = 0$ – the classic result for the predicted invasion speed $c_{mi} = 2\sqrt{D_i r_i}$ will hold, where c_{mi} is the monomorphic speed of strain i [53, 47]. From here on, we shall only discuss the case when all strains coexist – which requires that at least some of the mutation rates are non-zero.

We are interested in the case when all strains coexist and invade empty habitat, therefore we shall examine only those two hyperbolic equilibrium points: $\mathbf{n}^*_{trivial} = \mathbf{0}$ and $\mathbf{n}^*_{non-trivial} = (n^*_1, n^*_2, ..., n^*_N)$, where $\mathbf{n}^*_\iota \in \mathbb{R}^N_+$, $\iota \in \{trivial, non\text{-}trivial\}$, where subscript + denotes positive solutions. Note that the general system of spatially uniform equations has more than the aforementioned two equilibrium points which we do not examine. We are also only interested when these solutions are positive since they are associated to population densities – if these were negative then we would have negative population densitites, which have no biological relevance. We will follow the wave propagation technique described in van Saarloos [186] to calculate analytical predictions for the invasion speed. First we linearise equns. (2.1) around the unstable equilibrium point $\mathbf{n}^*_{trivial} = \mathbf{0}$, which leads to the form

$$\frac{\partial n_i}{\partial t} = D_i \frac{\partial^2 n_i}{\partial x^2} + \tilde{r}_i n_i + \sum_{j \in S \setminus \{i\}} \mu_{ji} n_j, \quad \forall i \in S, \tag{2.2}$$

where $\tilde{r}_i = r_i - \sum_{j \in S \setminus \{i\}}^{N} \mu_{ij}$. Note that the linearised system contains the parameters for dispersal, fecundity and mutation, but not intraspecific competition. Note also that the linearised system is co-operative, i.e. all strains have a positive effect on all other strains, while the non-linear system is strictly non-cooperative, since in equn. (2.1) intraspecific competition negatively impacts the other phe-

notypic populations. We can then substitute the ansatz

$$n_\iota \propto \exp(-i(\omega(k)t + kx)), \quad k, \omega(k) \in \mathbb{C}, \quad \forall \iota \in S,$$

into (2.2), where above $i = \sqrt{-1}$, and $\omega(k)$ and k are the dispersion relation and wave number respectively. It should be noted that, for biologically relevant solutions, k must be purely imaginary — otherwise solutions close to 0 would oscillate, producing negative population sizes. We then make the substitution $\lambda(q) = i\omega(k)$ and $k = iq$ where $q, \lambda(q) \in \mathbb{R}$, which will simplify the algebra. If we then drop the tildes from $\tilde{r}_i$ and make the simplifying assumption that $\mu_{ij} = \mu$, $\forall i \in S$ and $\forall j \in S \backslash \{i\}$, we have the following matrix:

$$\mathrm{M} = \begin{bmatrix} D_1 q^2 + r_1 - \lambda(q) & \mu & \cdots & \mu \\ \mu & D_2 q^2 + r_2 - \lambda(q) & \cdots & \mu \\ \vdots & \vdots & \ddots & \vdots \\ \mu & \mu & \cdots & D_N q^2 + r_N - \lambda(q) \end{bmatrix} \in \mathbb{R}^{N \times N}.$$

This then reduces to an eigenvalue problem which must solve the equation

$$\mathrm{M}\mathbf{N} = \mathbf{0},$$

where $\mathbf{N} \in \mathbb{R}^N$ is an eigenvector of the system. The characteristic equation is of the form

$$\prod_{i \in S} \left(D_i q^2 + r_i - \lambda(q) \right) = \mathcal{O}(\mu^2). \tag{2.3}$$

Since the RHS of equn. (2.3) approaches zero as $\mu \to 0$, we deduce that at least one of the terms within the product must be small when μ is small, therefore, without loss of generality we let

$$\varepsilon = D_l q^2 + r_l - \lambda(q) \quad \text{and} \quad \lim_{\mu \to 0} \varepsilon = 0, \quad l \in S, \tag{2.4}$$

where $\varepsilon \ll 1$ for some fixed l. We can then substitute equn. (2.4) into equn.

(2.3) to give

$$\varepsilon \prod_{i \in S \backslash \{l\}} \left(q^2 \psi_{il} - \varphi_{il} + \varepsilon \right) = \mathcal{O}(\mu^2) \tag{2.5}$$

where $\psi_{il} := D_i - D_l$ and $\varphi_{il} := r_l - r_i$ for $l \in S$ and $i \in S \backslash \{l\}$. Upon inspection of equn. (2.5), we can see that this equation can be expanded and rearranged into the form

$$\varepsilon = \frac{\mathcal{O}(\mu^2)}{\varepsilon^{N-1} + \varepsilon^{N-2} f_1(q) + ... + \varepsilon f_{N-2}(q) + f_{N-1}(q)}, \tag{2.6}$$

where $f_p(q)$ with $p = 1, 2, ..., N-1$ is some function comprised of the ψ_{il}, φ_{il}, and q terms from the product. Since we have assumed that $\varepsilon \ll 1$, equn. (2.6) is approximately given by

$$\varepsilon \approx \frac{\mathcal{O}(\mu^2)}{f_{N-1}(q)}, \tag{2.7}$$

which implies that $\varepsilon \to 0$ as $\mu \to 0 \iff f_{N-1}(q) \nrightarrow 0$, and we will retrieve the approximate invasion speeds given by Fisher and Kolomogorov et al. [53, 99]. However, if $f_{N-1} \to 0$ as $\mu \to 0$, then we must investigate further. To do so, we shall return to equn. 2.5 and use the method provided by van Saarloos [186], where the invasion speed can be determined by the critical point at which the phase velocity – the rate at which the phase of the wave travels through space – and the group velocity – the velocity at which the amplitude of the wave travels – of the wave are at parity. This is a standard technique for determining spreading rates of homogeneously stable states into unstable states, which is a reasonable assumption for an invasion. This is an appropriate technique to use since any invasion of this type will have a group velocity and phase velocity at parity i.e. one wave with a single amplitude. This relation is given by

$$\left.\frac{d\omega(k)}{dk}\right|_{k^*} = \frac{\Im\omega(k^*)}{\Im k^*} \quad \implies \left.\frac{d\lambda(q)}{dq}\right|_{q^*} = \frac{\lambda(q^*)}{q^*}, \tag{2.8}$$

where $\Im$ denotes imaginary parts. Further details can be seen in [186]. We first

take a derivative of equn. (2.5) with respect to q and rearrange for $\frac{\partial \varepsilon}{\partial q}$, to give

$$\frac{\partial \varepsilon}{\partial q} = \frac{-2\varepsilon q\left[\sum\limits_{i\in S\setminus\{l\}} \psi_{il}\left(\prod\limits_{j\in S\setminus\{l,i\}}\left(q^2\psi_{jl}-\varphi_{jl}+\varepsilon\right)\right)\right]}{\prod\limits_{i\in S\setminus\{l\}}\left(q^2\psi_{il}-\varphi_{il}+\varepsilon\right)+\varepsilon\sum\limits_{i\in S\setminus\{l\}}\left(\prod\limits_{j\in S\setminus\{l,i\}}\left(q^2\psi_{jl}-\varphi_{jl}+\varepsilon\right)\right)}. \tag{2.9}$$

We then substitute equn. (2.9) and equn. (2.4) into equn. (2.8) and take the appropriate derivative with respect to q which gives

$$\begin{aligned}\left(q^2D_l-r_l+\varepsilon\right)&\left[\prod_{i\in S\setminus\{l\}}\left(q^2\psi_{il}-\varphi_{il}+\varepsilon\right)...\right.\\ &\left....+\varepsilon\sum_{i\in S\setminus\{l\}}\left(\prod_{j\in S\setminus\{l,i\}}\left(q^2\psi_{jl}-\varphi_{jl}+\varepsilon\right)\right)\right]\\ &=-2\varepsilon q^2\left[\sum_{i\in S\setminus\{l\}}\psi_{il}\left(\prod_{j\in S\setminus\{l,i\}}\left(q^2\psi_{jl}-\varphi_{jl}+\varepsilon\right)\right)\right].\end{aligned} \tag{2.10}$$

Since we are interested in the behaviour when $\varepsilon \ll 1$ we take the limit as $\mu \to 0$ which gives

$$\lim_{\mu\to 0}\left(q^2D_l-r_l\right)\prod_{i\in S\setminus\{l\}}\left(q^2\psi_{il}-\varphi_{il}\right)=0.$$

We then have that either,

$$\lim_{\mu\to 0} q^2=\frac{r_l}{D_l},\quad \text{or}\quad \lim_{\mu\to 0} q^2=\frac{\varphi_{il}}{\psi_{il}}=\frac{r_l-r_i}{D_i-D_l},, \tag{2.11}$$

for $l \in S$ and $i \in S\setminus\{l\}$, where q is the spatial frequency used to define the wave speed when $\mu \to 0$. We then substitute equns. (2.11) into equn. (2.8) and find that the possible speeds of invasion are

$$\lim_{\mu\to 0} c_{ml}=2\sqrt{r_lD_l}, \tag{2.12}$$

or

$$\lim_{\mu\to 0} c_{dli}=\frac{|r_lD_i-r_iD_l|}{\sqrt{(r_l-r_i)(D_i-D_l)}}, \tag{2.13}$$

for $l \in S$, and $\forall i \in S\setminus\{l\}$, where c_{ml} is the monomorphic speed of any strain, and c_{dli} is the dimorphic speed given by any 2 strains – hereby referred to as the

"vanguard strains". Note that for $c_{dli} \in \mathbb{R}$, we require that $\big(r_l - r_i\big)\big(D_i - D_l\big) > 0$ is satisfied, which is the same result as found by Elliott and Cornell [47]. This means that the strain with the higher population growth rate out of l and i must have the lower dispersal ability, i.e. there must be a trade-off between dispersal and fitness. An important observation is that all other strains within the population do not contribute towards the predicted invasion speed.

In order for the above solutions for q and c_{dli} to represent the true behaviour of the system as $\mu \to 0$, we need to verify that both q and c_{dli} are real when μ is small but non-zero. This must be verified to ensure that around the steady state — $\mathbf{n}^*_{trivial} = \mathbf{0}$ — there are no oscillations, otherwise the population would become negative, and not biologically consistent. To do this, we examine the neighbourhood around $\mu = 0$ using a perturbative approach. Let us now fix l again as before, and also m, such that we have a fixed value for q. We can now perturb q^2 from equn. (2.11) by adding a small quantity that is dependent on μ, $\delta(\mu)$, to give

$$q^2 = \frac{\varphi_{ml}}{\psi_{ml}} + \delta(\mu), \quad l \in S, \quad m \in S\backslash\{l\}, \tag{2.14}$$

where $\psi_{ml} := D_m - D_l$ and $\varphi_{ml} := r_l - r_m$, and $\lim_{\mu\to 0}\delta(\mu) = 0$. We then substitute equn. (2.14) into equn. (2.10) and divide by ε to give the expression

$$\begin{aligned}\left(\eta + \delta D_l + \varepsilon\right)\Bigg[\left(\frac{\delta\psi_{ml}}{\varepsilon} + 1\right)\prod_{i\in S\backslash\{l,m\}}\left(\xi_i + \delta\psi_{il} + \varepsilon\right)... \\ ... + \sum_{i\in S\backslash\{l\}}\left(\prod_{j\in S\backslash\{l,i\}}\left(\xi_j + \delta\psi_{jl} + \varepsilon\right)\right)\Bigg] \\ = -2\frac{\varphi_{ml}}{\psi_{ml}}\left(\sum_{i\in S\backslash\{l\}}\psi_{il}\prod_{j\in S\backslash\{l,i\}}\Big(\xi_j + \delta\psi_{jl} + \varepsilon\Big)\right)... \\ ... - 2\delta\left(\sum_{i\in S\backslash\{l\}}\psi_{il}\prod_{j\in S\backslash\{l,i\}}\Big(\xi_j + \delta\psi_{jl} + \varepsilon\Big)\right),\end{aligned} \tag{2.15}$$

where $\eta := \frac{D_l\varphi_{ml} - r_l\psi_{ml}}{\psi_{ml}}$ and $\xi_i := \frac{\varphi_{ml}\psi_{il} - \varphi_{il}\psi_{ml}}{\psi_{ml}}$, $\forall i \in S\backslash\{m,l\}$. We should also

note, generally, that

$$\xi_i + \delta\psi_{il} + \varepsilon = \begin{cases} \varepsilon, & \text{for } i = l \\ \delta\psi_{ml} + \varepsilon, & \text{for } i = m, \\ \xi_i + \delta\psi_{il} + \varepsilon, & \text{otherwise} \end{cases} \qquad \forall i \in S. \tag{2.16}$$

We then take the limit when $\mu \to 0$ to give

$$\begin{aligned} \eta\left[\left(\psi_{ml}\lim_{\mu\to 0}\left(\frac{\delta}{\varepsilon}\right)+1\right)\prod_{i\in S\setminus\{l,m\}}\xi_i + \prod_{i\in S\setminus\{l,m\}}\xi_i\right] &= -2\varphi_{m1}\prod_{i\in S\setminus\{l,m\}}\xi_i \\ \iff \psi_{ml}\lim_{\mu\to 0}\left(\frac{\delta}{\varepsilon}\right)+2 &= -2\frac{\varphi_{ml}}{\eta}. \end{aligned} \tag{2.17}$$

From here we can see that $\frac{\delta}{\varepsilon} \to \mathcal{O}(1)$ as $\mu \to 0$, which in turn, implies that $\delta = \mathcal{O}(\varepsilon)$. To determine the conditions under which q is real we substitute equn. (2.14) into equn. (2.5) to give

$$\begin{aligned} \prod_{i\in S}\Big(\xi_i + \delta\psi_{il} + \varepsilon\Big) = -\mu^2\Bigg[\prod_{i\in S}\Big(\xi_i + \delta\psi_{il} + \varepsilon\Big)\sum_{\substack{j,k\in S\\ k>j}}\frac{1}{\Big(\xi_j + \delta\psi_{jl} + \varepsilon\Big)\Big(\xi_k + \delta\psi_{kl} + \varepsilon\Big)}\Bigg]... \\ ... + \mathcal{O}(\mu^3). \end{aligned} \tag{2.18}$$

We can then divide by μ^2 to give

$$\begin{aligned} \frac{\varepsilon^2}{\mu^2}\left(\frac{\delta\psi_{ml}}{\varepsilon}+1\right)\prod_{i\in S\setminus\{l,m\}}\Big(\xi_i + \delta\psi_{il} + \varepsilon\Big) = ... \\ ...\prod_{i\in S}\Big(\xi_i + \delta\psi_{il} + \varepsilon\Big)\sum_{\substack{j,k\in S\\ k>j}}\frac{1}{\Big(\xi_j + \delta\psi_{jl} + \varepsilon\Big)\Big(\xi_k + \delta\psi_{kl} + \varepsilon\Big)} + \mathcal{O}(\mu). \end{aligned} \tag{2.19}$$

Notice that many of the terms above include δ, ε, or μ, which then implies that in the limit $\mu \to 0$ we have

$$\lim_{\mu\to 0}\frac{\varepsilon^2}{\mu^2}\left(\frac{\delta\psi_{m1}}{\varepsilon}+1\right)\left(\prod_{i\in S\setminus\{l,m\}}\xi_i\right) = \prod_{i\in S\setminus\{l,m\}}\xi_i. \tag{2.20}$$

We then substitute equn. (2.17) into the above and rearrange to have

$$\lim_{\mu\to 0}\frac{\varepsilon}{\mu} = \pm\left[-\frac{\eta}{2\varphi_{m1}+\eta}\right]^{\frac{1}{2}}, \tag{2.21}$$

which informs us that $\frac{\varepsilon}{\mu} = \mathcal{O}(1)$ in the limit as $\mu \to 0$, which in turn, implies that $\varepsilon = \mathcal{O}(\mu)$ and $\delta = \mathcal{O}(\mu)$. We can then express this in terms of r_l, r_m and D_l, D_m as

$$\lim_{\mu\to 0}\frac{\varepsilon}{\mu} = \pm\left[\frac{r_l D_m + r_m D_l - 2r_l D_l}{r_l D_m + r_m D_l - 2r_m D_m}\right]^{\frac{1}{2}}. \tag{2.22}$$

Notice again that this condition is dependent on the traits of only two of the constituent strains within the population. For biologically relevant solutions, that is, q to be real, we require $\frac{\varepsilon}{\mu} > 0$ which implies that either

$$\frac{r_l}{r_m} + \frac{D_l}{D_m} \leqslant 2 \quad \text{and} \quad \frac{r_m}{r_l} + \frac{D_m}{D_l} \leqslant 2 \tag{2.23}$$

or

$$\frac{r_l}{r_m} + \frac{D_l}{D_m} \geqslant 2 \quad \text{and} \quad \frac{r_m}{r_l} + \frac{D_m}{D_l} \geqslant 2 \tag{2.24}$$

must be satisfied. However, since $r_i \geqslant 0$, and $D_i \geqslant 0$, $\forall i \in S$, then (2.23) can never be satisfied. If (2.24) is met then $\big(r_l - r_m\big)\big(D_m - D_l\big) > 0$ is automatically satisfied and we have a stronger condition for q, and likewise c_{dlm}, to be real. This is the same form of the conditions required for q and c_{dlm} to be real as was found for the two strain case by Elliott and Cornell [47]. The conditions for c_{dlm} require that both the dispersal and the growth rates for the two strains differ sufficiently – a topic we shall later address in section 2.4.

2.3 Numerical experiments

In the previous section we predicted analytical invasion speeds, but have no mathematical proof that the invasion will advance at any of the predicted speeds. Therefore, we use numerical experiments to simulate the invasion and to determine if the observed speed is indeed any of those which we have predicted. Experiments were conducted in MATLAB 2014a [115] using a finite difference θ-

method with $\theta = \frac{1}{2}$ — otherwise known as the Crank-Nicolson scheme [34]. The invasion speeds were calculated by recording 100 time points over the total temporal period and noting the spatial coordinates of the midpoint of the population density invasion front. The speed between each temporal point was calculated — speed = $\frac{\text{distance}}{\text{time}}$ — and an average of those speeds were taken. The results were compared with the analytic predictions and were found to be accurate within the error bounds of the numerical scheme and discretisation steps.

To test our theory we shall use a species that consists of 5 strains. We have selected r and D parameters such that each strain invades at the same monomorphic speed, but some of the dimorphic speeds are faster. Since our linear predictions for invasion speeds did not involve intraspecific competition coefficients we shall let these be neutral and equal. Finally, our solutions are valid in the limit as $\mu \to 0$, therefore we shall allow $\mu_{ij} = \mu_{ji} = 0.001$. The parameters for simulations are given in table 2.1 unless otherwise stated.

Table 2.1: General parameter values used within numerical simulations given in figure 2.1

Descriptor	**Symbol**	**Values**
Spatial step size	Δx	0.1
Temporal step size	Δt	0.1
Total landscape size	L	[0,35000]
Total temporal period	T	10,000
Dispersal coefficients	$(D_1, D_2, D_3, D_4, D_5)$	$(8.5, 8, 4.5, 3, 0.5)$
Per capita growth rate	$(r_1, r_2, r_3, r_4, r_5)$	$(0.05, 0.3, 0.45, 0.8, 0.85)$
Mutation rate	μ	0.001
Intraspecific competition	$C_{ij} = C_{ji} = \frac{1}{K}$	1
Boundary conditions	∂L	$\frac{\partial \nu(T,L)}{\partial x} = 0, \nu = \{v_i\}$
Initial conditions	$v_i = [0, 3500]$	$(\frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5})$

2.3.1 5 strain species

Our numerical experiments have shown that a species with 5 strains invades at the same predicted speed as that containing only the 2 that contribute towards the dimorphic speed in equn. (2.12). Fig. 2.1 panel a gives an example of a species consisting of 5 strains, where the invasion speed is higher than any of the monomorphic speeds but the same as for a species containing only strains 2 and 4 (panel b), which are the strains whose traits give the highest value of the dimorphic speed c_{dij} in equn. (2.13). Thus, while an anomalous invasion implies a synergy between two strains (benefiting from the growth rate of a more fecund strain and the dispersal ability of a more dispersive one), this synergy does not extend beyond more than two strains. Moreover, one might expect that the vanguard strains would be the most dispersive and the most fecund, so that the population as a whole benefits from the highest diffusion constant and the highest population growth rate. However, it turns out that this need not be the case. In the example shown in fig. 2.1, the fastest dimorphic speed c_{dij} – and thus the invasion speed for the 5-strain species – is that for strains 2 and 4 (panel b), whereas strain 1 is the most dispersive, and strain 5 has the highest population growth rate, where both together (panel c) are slower than strains 2 and 4. A summary of the predicted and calculated speeds can be seen in table 2.2.

Table 2.2: Invasions speeds predicted from analytical solutions and calculated from numerical experiments associated with figure 2.1

Strains within population	**Analytically predicted speed**	**Observed invasion speed**
5 strains	3.478	3.345
strains 2 & 4	3.478	3.387
strains 1 & 5	2.841	2.7825
strain 2 in isolation	3.09	3.051
5 strains mutation only between neighbouring strains	3.478	3.366

2.3.2 Mutation between neighbouring strains

We have found that invasion speeds do not depend on the relative values of the mutation rates, even if some mutation rates are zero (provided the system does not factorise into independent subsets of strains between which mutation is impossible). In particular, the invasion speed is the same for a system with "universal" mutation ($\mu_{ij} = \mu$ for all i and j) as for "nearest neighbour" mutation ($\mu_{ij} = 0$ unless $|i - j| = 1$) (fig. 2.1 e). This means that our results should not only apply to species with a small set of discrete strains but also extend to the case of a very large, effectively continuous, set of strains where the traits can only mutate by small amounts at each generation.

2.3.3 Non-neutral interactions

We have seen that our analysis predicts that the invasion speed is determined by the dynamics at low densities, and is therefore independent of the competition coefficients C_{ji}. We should, therefore, obtain the same anomalous speeds, whether or not the strains coexist within the range core – i.e. at the stable equilibrium. Parameter values for these simulations are shown in table 2.2 and results are illustrated in fig. 2.2. The interaction matrix for these sets of experiments was given by

$$C = \begin{bmatrix} 1 & 1.25 & 0.8 \\ 0.8 & 1 & 0.8 \\ 0.8 & 0.8 & 1.25 \end{bmatrix}. \tag{2.25}$$

The interaction matrix thus represents a three-strain species, where strain 2 out-competes both strains 1 and 2 in the range core. fig. 2.2 shows that the invasion speed of a three-strain species (panel a) is given by the anomalous speed for strains 1 and 3 (panel b), even though these are both out-competed by strain 2 in the range core. This shows that strains that are very rare in the range core of the species can still determine the invasion dynamics of the species.

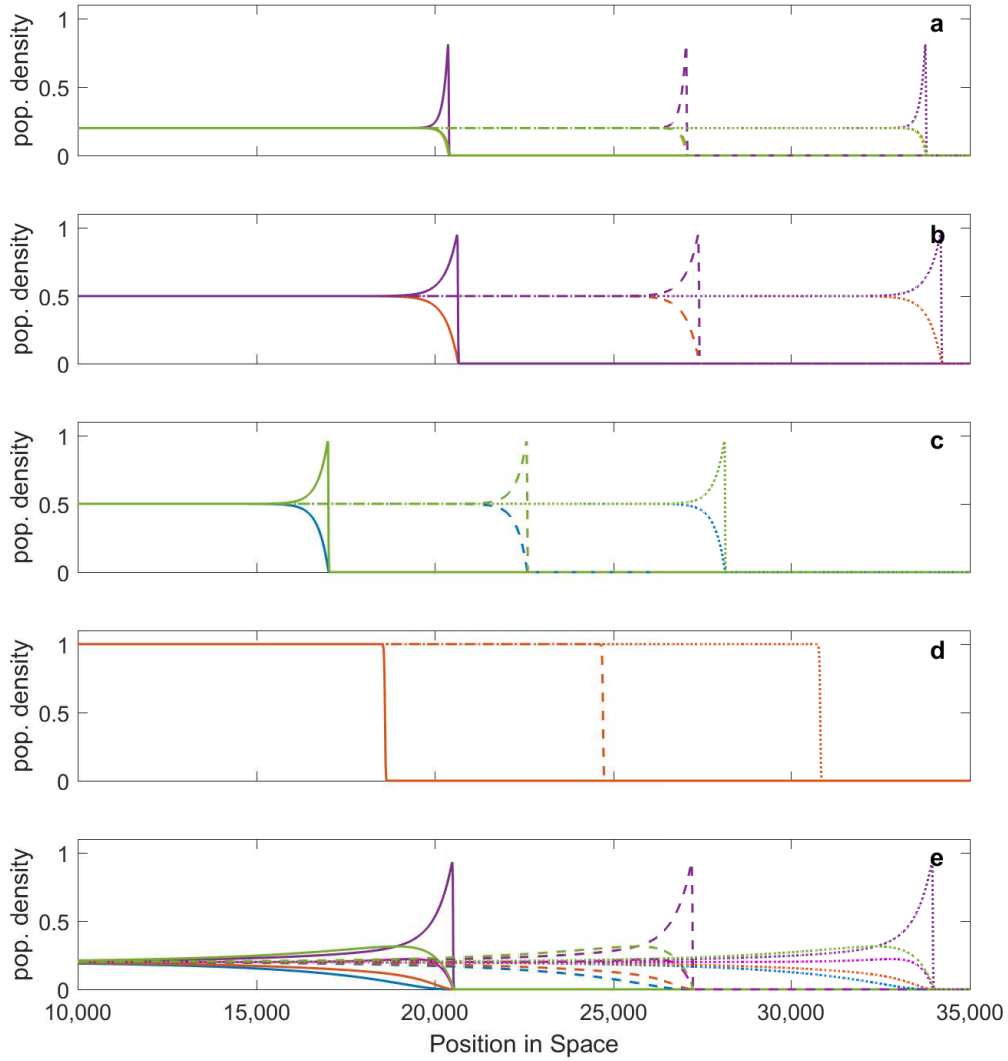

Figure 2.1: 5 different cases of polymorphic invasion into unoccupied habitat. Panel a: all 5 strains are present within the population, with equal mutation rates between all strains. Panel b: species contains only the two vanguard strains from panel a (strains 2 and 4), giving the same invasion speed as the 5-strain species in panel a. Panel c: 2-strain species consisting of the most dispersive and most fecund strains (strains 1 and 5). Invasion speed is anomalous (faster than either strain in isolation, not shown) but slower than for panels a and b. Panel d: species contains only the strain with the fastest monomorphic speed (strain 2), which is slower than cases a and b. Panel e: same 5 strains as in Panel a, but with mutation only between neighbouring strains ($|i-j|=1$), showing that the invasion speed is the same as for universal mutation as shown in a. Parameters: Strain 1 (blue) $(D,r)=(8.5,0.05)$; Strain 2 (orange) $(D,r)=(8,0.3)$; Strain 3 (yellow) $(D,r)=(4.5,0.45)$; Strain 4 (purple) $(D,r)=(3,0.8)$; Strain 5 (green) $(D,r)=(0.5,0.85)$. Time snapshots taken at $t=6000$ (solid lines); $t=8000$ (dashed lines); $t=10,000$ (dotted lines). Simulations were run with other parameters set as $C_{ij}=1$, $\mu_{ij}=0.001$ or zero depending on whether mutation is universal or nearest neighbour.

Table 2.3: General parameter values used within numerical simulations given in figure 2.2

Descriptor	**Symbol**	**Values**
Spatial step size	Δx	0.1
Temporal step size	Δt	0.1
Total landscape size	L	[0,35000]
Total temporal period	T	10,000
Dispersal coefficients	(D_1, D_2, D_3)	$(8.5, 4.5, 3)$
Per capita growth rate	(r_1, r_2, r_3)	$(0.05, 0.45, 0.8)$
Mutation rate	μ	0.001
Boundary conditions	∂L	$\frac{\partial \nu(T,L)}{\partial x} = 0, \nu = \{v_i\}$
Initial conditions	$v_i = [0, 3500]$	$(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$

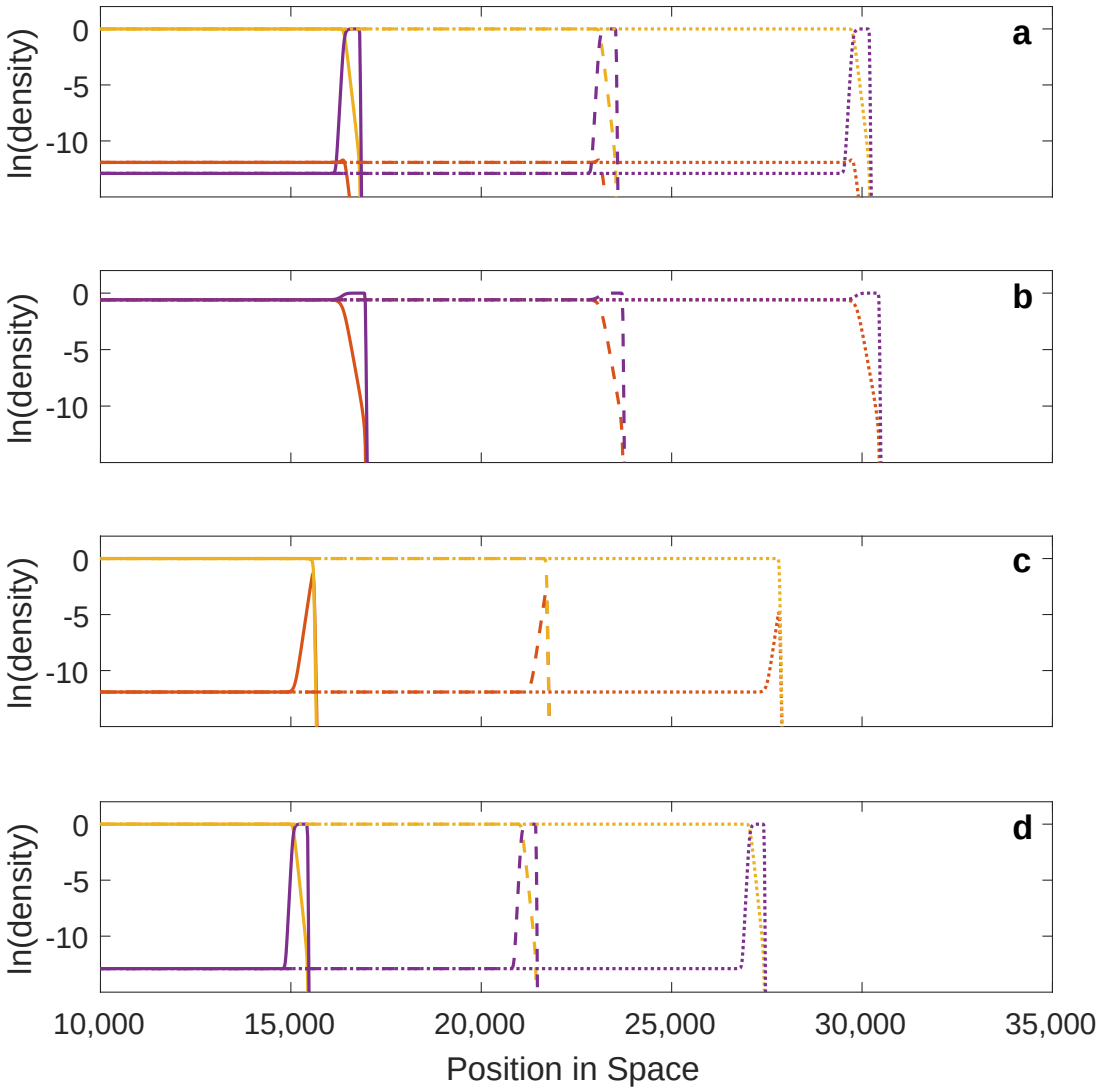

Figure 2.2: Strains that are outcompeted at equilibrium can still generate anomalous invasion speeds. Panel a: Invasion by 3-strain species. Strain 3 dominates the core while strains 2 and 4 are present in low numbers, but strain 4 leads the invasion. Panel b: Species consisting only of strains 2 and 4 invades at the same anomalous speed as the 3-strain species shown in panel a. Species consisting only of strains 2 and 3 (Panel c) or strains 3 and 4 (Panel d) invade more slowly than in panel a and b. Parameters: Strain 2 (orange) $(D, r) = (8, 0.3)$, Strain 3 (yellow) $(D, r) = (4.5, 0.45)$, Strain 4 (purple) $(D, r) = (3, 0.8)$. Time snapshots taken at t=4,000 (solid lines); t=6,000 (dashed lines); t=8,000 (dotted lines). Simulations were run with a competition matrix C_{ji} = [1, 1.25 0.8; 0.8, 1, 0.8; 0.8, 1.25, 1], $\mu_{ij} = 0.001 \forall i, j$. Log density is shown so that low population densities are visible.

2.4 Invasion speed determined by r-D trade-off

We have seen that many dimorphic speeds are possible, and our previous results have alluded that the values of both r and D are important. In nature any investment in dispersal must be accompanied by a trade-off with another trait. Trait values are continuous in nature, therefore we can then construct trade-off curves in $r - D$ space to illustrate the relationship. Since we assume a $r - D$ trade-off, we should expect that the structure of such a curve should determine the invasion speed, we then construct some hypothetical curves that illustrate the a selection of scenarios. This allows us to predict the invasion speed for a species with a continuous set of strains.

2.4.1 Monomorphic speed

To predict a monomorphic speed, we assume that each strain has a unique phenotype determined by r and D, and that there is a trade-off between r and D so that $r(D)$ is a decreasing function of D. The invasion speed, after a sufficiently long time, will again be given by the largest permitted value of c_{mi} and c_{dij} calculated using equns. (2.12, 2.13, 2.24) for all strains, and pairs of strains of the species. It turns out that these equations have a geometric interpretation and that the invasion speed depends, in a simple way, on the shape of the $r(D)$ curve. In particular, the existence of an anomalous invasion speed is determined by the curvature of the trade-off curve.

We first consider the possible monomorphic speeds for this species. Eqn (2.12) shows that the strain with the highest value of $r(D)D$ has the fastest monomorphic speed. fig. 2.3 illustrates the range of scenarios. If $r(D)D$ diverged at high or low D, then we would expect the invasion speed to increase without limit. This could imply that D evolves to ever higher values (fig. 2.3 a dotted curve) or ever lower values (fig. 2.3 b, solid curve). However, for real species any strain will have finite values for $r(D)$ and D, so there will be a maximal value for these traits. Depending on the details of the relationship, the fastest monomorphic speed could be the one where D takes the highest (fig. 2.3 a,c) or

lowest (fig. 2.3 b) of its possible values, or an intermediate value (fig. 2.3 d). This will only be the ultimate invasion speed for the species, however, if it is greater than the fastest valid dimorphic speed.

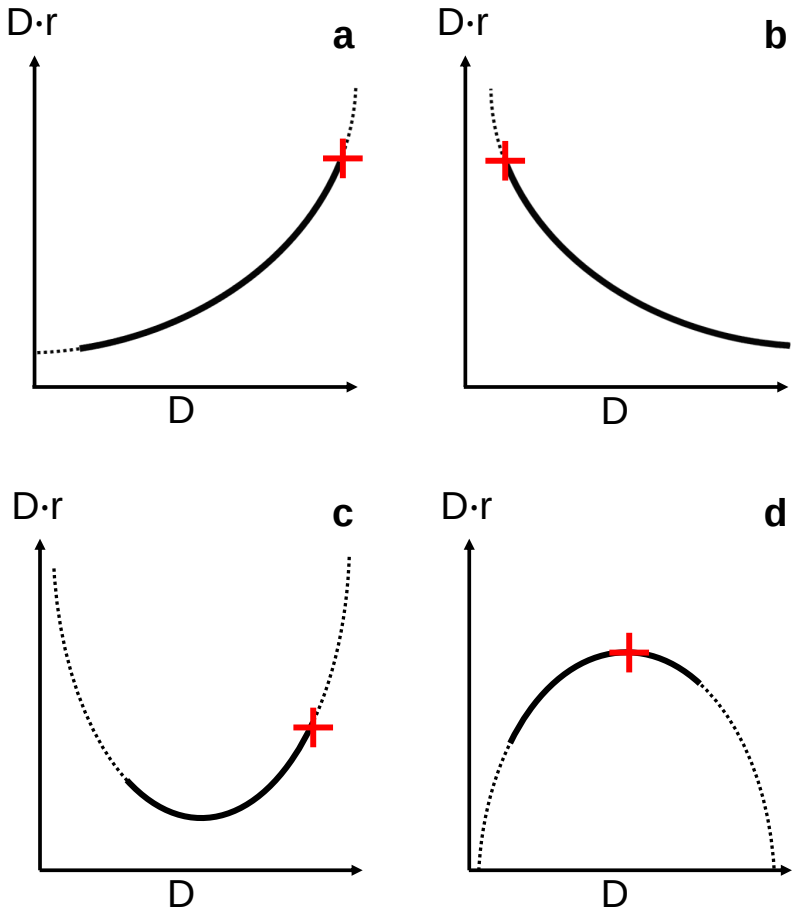

Figure 2.3: The shape of the curve of $D \times r(D)$ as a function of D determines the fastest monomorphic invasion speed, and therefore the asymptotic invasion speed when there are no faster anomalous speeds. If the trade-off is such that $D \times r(D)$ could grow without limit, either at high D (Panel a dotted curve) or low D (Panel b dotted curve), then the invasion would accelerate without limit. If $r(D)$ and D are restricted to a finite range of values (solid curves), the invasion speed approaches an asymptotic value given by the strain with the highest value of $D \times r(D)$ (red crosses). In many cases, particularly when the $D \times r(D)$ curve has positive curvature, this will be the strain with either the highest (Panels a, c) or the lowest (Panel b) values of D. However, if the $D \times r(D)$ curve has negative curvature then the fastest speed could be that for a strain with an intermediate value of D.

2.4.2 Dimorphic speed

Turning now to dimorphic speeds, we find the discussion is simplified by a geometric interpretation (see Appendix A) of equn (2.13), illustrated in fig. 2.4. The dimorphic speed for two strains is equal to the monomorphic speed for a "virtual" strain, which lies at the midpoint (green triangles in fig. 2.4) of the straight-line segment joining the two axes – dotted line in fig. 2.4 – and passing through the points representing the two strains – blue circles in fig. 2.4 – in $r(D)$-D space. If the virtual strain lies between the two real strains, then conditions (2.24) are met and this is a valid dimorphic speed (figs. 2.4 a,c,e,f); otherwise, this is not a valid dimorphic speed for the species.

Therefore, if the trade-off curve is a straight line (fig. 2.4 a,b), then all pairs of strains have the same virtual strain – and therefore the same dimorphic speed – whether or not the trade-off curve encompasses this virtual strain, when the fastest possible speed will be the fastest monomorphic speed among the constituent strains. On the other hand, if the trade-off curve has positive curvature, fig. (2.4 e), then the virtual strain for any pair of strains either lies below the trade-off curve, or does not lie between the two real strains. In that case, none of the valid anomalous speeds are faster than the fastest monomorphic speed for the species. In both of these cases, the ultimate invasion speed will be the same as the monomorphic speed for the fastest strain in isolation.

However, if the trade-off curve has positive curvature (fig. 2.4 c,d), then the line segment joining the two most extreme strains will lie above the trade-off curve. If the range of $r(D)$ and D values is wide enough that the corresponding virtual strain lies between these extreme strains (fig. 2.4 c), then this will have the fastest dimorphic speed, which will be faster than any of the constituent monomorphic speeds. However, if the range of $r(D)$ and D values are not wide enough (fig. 2.4 d), the species will invade at the fastest monomorphic speed. In the former case, the invasion will follow the anomalous speed when the vanguard strains have the highest $r(D)$ and the highest D. A further possibility is that the curvature of the trade-off curve is positive in some places and negative in others

(fig. 2.4 f), for example if the trade-off is more acute at more extreme values . In this case, the vanguard strains will be the ones where the line connecting them is tangential to the trade-off at both points, which will not represent the most extreme traits in the species (similarly to what was found in fig. 2.1 for a species with a discrete set of trait values).

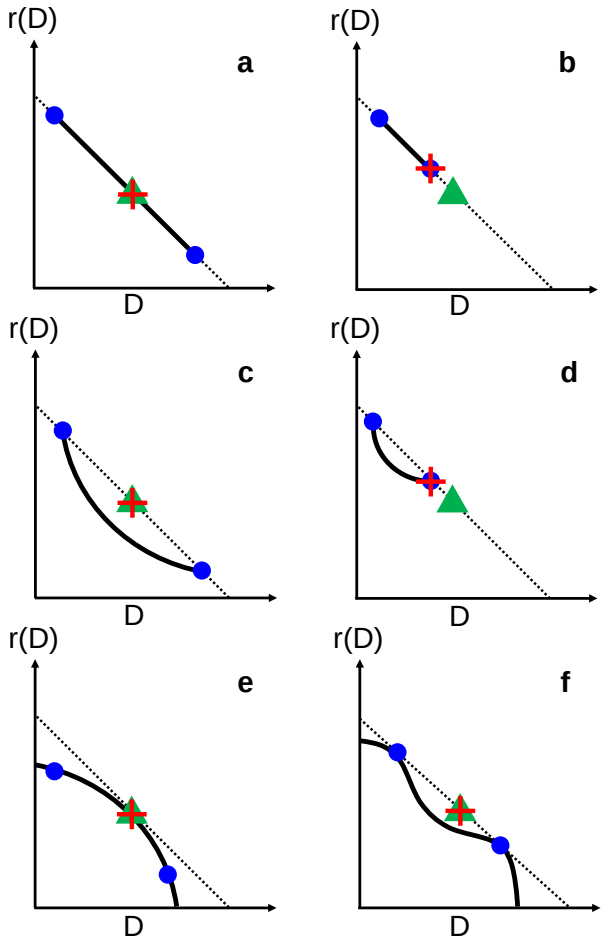

Figure 2.4: The geometry of the D-$r(D)$ trade-off determines which anomalous speeds, if any, give the eventual invasion speed for a species comprising a continuum of strains. Solid lines represent the trade-off curve between D and $r(D)$. Dotted lines are chords that pass through points on the trade-off curve and terminate at the axes. Green triangles are the midpoint of these chords, and represents the "virtual strain" for any two strains that the chord passes through (see section 2.4.2). Blue points represent two particular strains of the species. The highest invasion speed in each case species is given by $2\sqrt{D_+ r_+}$, where D_+ and r_+ are evaluated at the red crosses. Panels a,b: for a straight-line trade-off, all pairs of species have the same chord and therefore the same virtual strain. In Panel a the virtual strain is a constituent of the species, but not in Panel b. In Panel c the virtual strain for the two extreme strains (blue circles) lies between them, and since the trade-off has positive curvature the virtual strain has a faster speed than any single constituent. In Panels b and d the virtual strain does not lie between any pair of points on the trade-off curve, so this does not yield a valid anomalous invasion speed. If the trade-off has negative curvature (Panel e) there are valid anomalous speeds, but none of their virtual strains lie above the trade-off so the asymptotic invasion speed is the fastest constituent monomorphic speed. In Panel F the chord is tangential to the trade-off at the two blue points, and since the whole trade-off curve lies below this chord the corresponding anomalous speed is faster than any other monomorphic or dimorphic speed for this species. This shows that the vanguard strains are not necessarily the ones with the highest values of $r(D)$ or D. We show further in the Appendix A that the vanguard morphs do not necessarily have the highest value of $r(D)D$ either.

2.5 Discussion

We have presented a general theory for invasions by a species where dispersal ability varies among lineages and trades off against fitness. In many cases, invasions advance at the same rate as if the population was made up of the single strain that has the highest monomorphic speed. However, if the dispersal-growth curve has positive curvature, and extends over a wide enough range of dispersal rates, then an anomalous invasion speed occurs which is faster than that for any single constituent strain, but depends only on the traits of two "vanguard" strains. The vanguard strains need neither be the most dispersive nor the most fecund. If the growth or dispersal ability can take arbitrarily large values (which is unlikely for any real species) then the invasion front could accelerate continuously rather than approaching an asymptotic speed. The invasion speed is insensitive to the mutation rates between strains or the details of inter-strain competition, and anomalous speeds can even be caused by strains that are outcompeted at equilibrium (fig. 2.2). Previous simulation studies suggests that evolution at range fronts should select for the most dispersive strategies [181, 161, 135, 55], but we show that this need not be the case. For example, in fig. 2.1 the invasion is led by the most dispersive (at the very forefront of the wave) of the vanguard strains, which is not always the most dispersive of all the strains.

We noted in section 2.2 that a proof that the system described in equn. (2.1) is linearly determined. The main result of Girardin's [62] work is that the invasion speed of the system is determined by the linear prediction, and furthermore, the invasion speed is given by the Perron-Frobenius eigenvalue of the system – the eigenvalue associated with the fastest speed. The results of Girardin give our findings greater weight that anomalous invasion speeds are possible, and are determined by the shape of the $r - D$ trade-off curve.

It is important to note that anomalous speeds persist even when the mutation rates between strains are vanishingly small (i.e. in the limit $\mu_{ij} \to 0$). This is surprising because, when the mutation rates are strictly zero, anomalous speeds do not occur and the strains invade independently at their monomorphic speeds

– or not at all, if they are outcompeted in the stable equilibrium [47]. However, a small amount of mutation, combined with exponential growth, is sufficient for the different strains to keep up with each other during the invasion and participate in the invasion. Other systems with anomalous invasion speeds are characterised by strongly co-operative interactions between strains, i.e. each strain has a strongly positive effect on the other [193, 194]. In our system, mutation is the only co-operative interaction between strains, so our case is unusual because anomalous invasion speeds exist even when cooperation (i.e. mutation) is vanishingly weak.

Our results predict that invasion dynamics depend critically on the curvature of the trade-off curve between dispersal and population growth rate, but while the existence of such trade-offs has been established [86, 66, 91] their shape has not been quantified empirically in much detail. This trade-off arises from the organism diverting resources either to dispersal or to reproductive success, but the rates describing these abilities depend on the details of the organism's anatomy and physiology so the trade-off curve could in principle take many different shapes. One plausible assumption would be that the organism's reproductive success is proportional to the energy diverted to reproductive organs, and the distance of each dispersal event is proportional to the energy diverted to movement organs, which would imply a straight-line trade-off between fitness and dispersal distance. However, while population growth rate is directly proportional to fitness, the diffusion constant is proportional to the square of the dispersal distance, which would imply that the $D(r)$ curve would be a parabola with positive curvature. On the other hand, diminishing returns would imply that an incremental improvement in one trait comes at a much greater cost when that trait is at the higher end of its range of possible values, which would suggest a trade-off with negative curvature (e.g. fig. 2.4 e) or possibly a more complex curve such as in fig. 2.4 f. Anomalous speeds could also occur when the trade-off curve has positive curvature, provided the species exists in distinct morphs, such as wing dimorphic crickets, so that the trade-off is not a single continuous curve – and that the virtual strain with the fastest invasion speed lies between different segments of the trade-off curve. Thus, while anomalous speeds can be expected in a

wide range of scenarios, to predict the species in which they occur we would need more detailed measurements of the trade-offs between dispersal and reproductive fitness than are currently available – something we address in chapter 4.

We should note that the system of equations (2.2) has several Fourier modes that are unstable, and as such we have several q values that each have different linear spreading velocities. For the linearised system (2.2), the predicted invasion speed would be the fastest of these velocities [186], but it should be noted that this is not necessarily the case for the fully non-linear system (2.1). However, it has since come to our attention after producing this work that a proof has been derived that the linearly predicted speed is indeed the speed that the invasion will invade at [62].

Our model may be simple and generic, but we expect the predictions to apply in a wide range of scenarios. Similar anomalous invasion speeds are expected if space and/or time are discrete rather than continuous [48]. We used Brownian motion to model dispersal (i.e. assuming dispersal comprises many very short steps), but the analysis and our predictions would be similar if dispersal followed a jump process with thin-tailed (exponentially bounded) dispersal kernels. At first sight, our theory would appear not to apply to the case of dispersal with fat-tailed (i.e. not exponentially bounded) dispersal kernels, which have been used to describe dispersal combining local movement and long-distance events [25, 24, 100]. Such kernels have been shown to predict accelerating invasions rather than a travelling wave of constant speed, such as those in the Cane Toad invasion in North-Western Australia [136]. However, it has been shown that fat-tailed dispersal kernels can describe a population of individuals, each performing Brownian motion but with a distribution of diffusion constants [134], which is equivalent to the dispersal polymorphism we discuss in this paper. If the population growth rate is independent of the diffusion coefficient – as assumed in Petrovskii et al. [134] – the $r(D)$ trade-off curve will resemble fig. 2.3 a, and our theory does indeed predict accelerating invasions. However, if the more dispersive strains have a lower population growth rate, then our theory predicts that the

invasion could follow an asymptotic constant invasion speed determined by the shape of this trade-off. This shows that it could be misleading to characterise a species by a single dispersal kernel, without considering whether intrinsic dispersal ability or reproductive ability might differ among individuals. We expect a more complex theory is needed to account for other factors such as non-linear diffusion or landscape heterogeneity, but our results constitute the first steps in this direction.

Our model assumes a haploid species, but we predict our theory to hold in any species where offspring inherit their strain (with rare mutation) from only one parent when population densities are low. The theory should apply for microbes as well as self-fertile plants, but cannot be simply extended to obligate sexually reproducing species with recombination where the population dynamics at low densities are non-linear. Our theory is deterministic, but we expect it to describe species where local carrying capacities and/or mutation rates are large, such as in microbes and insects [48]. In particular, we expect it to apply to the dynamics of an infectious disease, where each host constitutes a deme with a large local population of the pathogen.

Chapter 3

Polymorphic invasions in the presence of a competitor

3.1 Introduction

In chapter 2 we explored how a polymorphic species can invade vacant favourable habitat. Since species are shifting their ranges, they will inevitably disperse into habitat where they will encounter other resident species with whom they must compete for resources. Such expansions could herald a loss of biodiversity by crashing native populations, threaten the water resources by displacing natives that retain more water [188, 114], or increase the fire hazard risk if invasives produce more flammable dry biomass then the native [16, 37]. Therefore, understanding how invasive species spread in the presence of competitors, and how quickly they do so, is an important area of study that could be used to minimise these risks.

Much of the past research of competitive effects during biological invasions has focused on native species resisting invasion from an exotic newcomer, but little attention has been given to the role that competition has on invasion speeds [49, 40, 162, 79, 111, 78]. Recent research has been primarily theoretical, with very

little empirical experiments due to the impracticalities of measuring competitive effects during invasion. Many of these theoretical studies have shown a general consensus that competition slows the rate of expansions [129, 108, 77]. This has been validated in some empirical systems such as the spread of cane tibouchina (*Tibouchina herbacea*) on the Big Island of Hawai'i was shown to invade slower when competing against forest communities [4], or native perennials slowing the advance of exotic annual grasses in California [158].

Despite these advances, almost all of the previous studies have assumed that the interacting species are monomorphic. When we considered the invasion of a polymorphic species in chapter 2 we found that it could invade anomalously, which suggests that assuming a species is monomorphic could lead to erroneous predictions. A natural question to ask is then: is this phenomenon robust when there is a competitor present? Results derived by Girardin that prove anomalous invasion speeds travel at those predicted from linearisation cannot be readily applied when a resident is interacting with the polymorph [62]. The proof requires that the eigenvalue problem is irreducible, whereas the introduction of a resident violates this requirement. We have seen before that intraspecific competition between strains had no effect on the invasion speed, but it is not obviously clear that this would also be the case for interspecific competition between species. Accordingly, it is unknown if polymorphism will aid an invading species, or help to resist encroachment if it is being invaded.

Therefore, we develop a theory that builds upon the previous chapter by incorporating competition with another species. We shall consider the simplest case that this scenario offers, which is two species in competition where one has two-strains (dimorphic) with a dispersal-reproduction trade-off, and the other has a single strain (monomorphic). Firstly, we shall find conditions under which each species can be invaded, followed by analytically deriving invasion speeds when either species can invade the other. Since our approach is not rigorous — but should provide a lower bound for the invasion speed [99] — we shall compare these predictions with speeds determined from numerical simulations to find if

both are in agreement. We shall find that either species can invade the other if the competitive effects of the other are weak enough, and that anomalous invasion speeds are possible for the dimorph. Furthermore, we will find that the system can be re-parametrised such that competition can be absorbed into the per capita growth rate to reduce the problem to that of chapter 2. We shall also observe that the effect of competition reduces the invasion speed. Finally, and perhaps most interestingly, we will show that anomalous invasion speeds are still possible when competition is strong enough to enforce negative growth on one of the dimorphic strains.

3.2 Mathematical analysis

To account for the inclusion of another population, let us form the model from the following assumptions.

- Each experiences non-neutral interspecific competition.
- Both species consist of asexual haploid individuals.
- The polymorphic population consists of strains which differ in their reproductive, dispersal, and interspecific competitive abilities.
- Progeny from each strain can mutate from their parental strain to ther within the species.
- Both species experience overlapping generations.
- The entirety of the habitat is favourable to both species.
- Time and space are both considered as continuous.

With our assumptions in place we can then derive the model from first principles as a set of equations given by,

$$\begin{aligned}\frac{\partial v_i}{\partial t} &= \overbrace{D_i \frac{\partial^2 v_i}{\partial x^2}}^{\text{dispersal}} + \overbrace{b_i v_i - d_i v_i}^{\text{births and deaths}} - \overbrace{\sum_{j=1}^{N} \delta_{ji} v_j v_i - \delta_{ui} u v_i}^{\text{competition between strains and species}} + \overbrace{\sum_{j=1}^{N} b_j \alpha_{ij} v_j - b_i v_i \sum_{j=1}^{N} \alpha_{ij}}^{\text{mutation dynamics}}, \\ \frac{\partial u}{\partial t} &= D_u \frac{\partial^2 u}{\partial x^2} + b_u u - d_u u - u \sum_{j=1}^{N} \delta_{iu} v_j - \delta_{uu} u,\end{aligned} \tag{3.1}$$

where our assumptions are represented as variables and parameters in the following way,

- v_i is polymorphic species density, and subscript i denotes the ith phenotype.
- u monomorphic species density.
- b_ι birth rate of species or strain ι.
- d_ι mortality rate by natural processes of species or strain ι.
- $\delta_{\kappa\iota}$ mortality of species or strain ι due to the presence of species or strain κ.
- α_{ij} probability of mutation at birth from strain i into strain j.

Note that subscript $\iota, \kappa \in \{1, 2, ..., N, u\}$ refers to each individual phenotypic strain or to the monomorphic species. Equn. (3.1) can then be re-parameterised into a form more easily analysed to give

$$\begin{aligned}\frac{\partial v_i}{\partial t} &= D_i \frac{\partial^2 v_i}{\partial x^2} + r_i v_i \left(1 - \sum_{j=1}^{N} C_{ji} v_j - C_{ui} u\right) + \sum_{j=1}^{N} \mu_{ji} v_j - v_i \sum_{j=1}^{N} \mu_{ij}, \\ \frac{\partial u}{\partial t} &= D_u \frac{\partial^2 u}{\partial x^2} + r_u u \left(1 - \sum_{i=1}^{N} C_{iu} v_i - C_{uu} u\right),\end{aligned} \tag{3.2}$$

where $r_i = b_i - d_i$ is the per-capita population growth rate; $C_{ji} = \frac{\delta_{ji}}{b_i - d_i}$ are the competition coefficients of subscript j on subscript i; and $\mu_{ij} = b_i \alpha_{ij}$ is mutation rate of phenotype i mutating at birth into phenotype j. In order to remain biologically consistent, we should introduce some conventions regarding the terms in the aforementioned equations.

- Variables representing the species densities will be strictly non-negative, i.e. u, and $v_i \geq 0$.

- Our parameters C_{ji} must be positive to represent competition, if they are negative this implies co-operation, i.e. $C_{ji} > 0$.

- Mutation rate, $\mu_{ij} > 0$

- Growth & dispersal rates are strictly non-negative, i.e. $r_i, D_i \geq 0$.

It is also given that none of the above terms can ever be non-real, as described in chapter 2.

The general form of the two species case is given by equn. (3.2), however, it can be extended to the multi-species case with the addition of more equations. To remove unnecessary complexity, we will restrict ourselves to the two species case, one of which is dimorphic, and the other monomorphic. At this point we will simplify the model such that each species competes neutrally with conspecifics; and by assuming that mutation between strains are equal, i.e. $\mu_{ij} = \mu_{ji} = \mu$. Note that this assumption has been made in chapter 2 where we seen that differences in mutation rate were unimportant, since the model is similar, we should expect

this to remain the case. The set of equations is then given by

$$\begin{aligned}
\frac{\partial v_1}{\partial t} &= D_1\frac{\partial^2 v_1}{\partial x^2} + r_1 v_1\Bigg(1 - v_1 - v_2 - C_{u1}u\Bigg) + \mu(v_2 - v_1),\\
\frac{\partial v_2}{\partial t} &= D_2\frac{\partial^2 v_2}{\partial x^2} + r_2 v_2\Bigg(1 - v_1 - v_2 - C_{u2}u\Bigg) + \mu(v_1 - v_2),\\
\frac{\partial u}{\partial t} &= D_u\frac{\partial^2 u}{\partial x^2} + r_u u\Bigg(1 - C_{1u}v_1 - C_{2u}v_2 - u\Bigg).
\end{aligned} \tag{3.3}$$

Since we are interested in the situations when an invading species supplants an already resident species from the habitat, we are only interested in two particular equilibria; when $\mathbf{N}^* = (v_1^*, v_2^*, u^*)$ is either, $\mathbf{N}^* = (\frac{1}{2}, \frac{1}{2}, 0)$, or $\mathbf{N}^* = (0, 0, 1)$. This allows to us to explore scenarios when a dimorph is invading a resident monomorph, and *vice versa*.

3.2.1 Stability of population equilibria

If we want to determine whether a species invades or not, we must first examine the stability of the equilibria. We can determine the stability of the equilibria densities by finding the eigenvalues of the Jacobian matrices of equn. (3.3) evaluated at these points, which are

$$J|_{(\frac{1}{2},\frac{1}{2},0)} = \begin{bmatrix} \frac{-r_1}{2} - \mu & \frac{-r_1}{2} + \mu & \frac{-r_1 C_{u1}}{2} \\ \frac{-r_2}{2} + \mu & \frac{-r_2}{2} - \mu & \frac{-r_2 C_{u2}}{2} \\ 0 & 0 & r_u\Bigg(1 - \frac{(C_{1u}+C_{2u})}{2}\Bigg) \end{bmatrix}, \tag{3.4}$$

and,

$$J|_{(0,0,1)} = \begin{bmatrix} r_1(1 - C_{u1}) - \mu & \mu & 0 \\ \mu & r_2(1 - C_{u2}) - \mu & 0 \\ -r_u C_{1u} & -r_u C_{2u} & -r_u \end{bmatrix}. \tag{3.5}$$

We shall now calculate the eigenvalues of equns. (3.4) and (3.5) and de-

termine under which conditions each of these equilibria are unstable. This is achieved by finding when the real parts of the eigenvalues are strictly positive, i.e. $\mathrm{Re}(\lambda) > 0$.

Let us begin with equn. (3.4), which when unstable, corresponds to the polymorphic species being susceptible to invasion. We now find when the determinant of the matrix in equn. (3.4) is equal to zero ($\det(J|_{(\frac{1}{2},\frac{1}{2},0)} - \lambda I) = 0$), which in turn provides us with the characteristic equation,

$$\mathrm{AB} = 0$$

where

$$\begin{aligned} A &= r_u\left(1 - \frac{(C_{1u} + C_{2u})}{2}\right) - \lambda, \\ B &= \left(\frac{-r_1}{2} - \mu - \lambda\right)\left(\frac{-r_2}{2} - \mu - \lambda\right) - \left(\frac{-r_1}{2} + \mu\right)\left(\frac{-r_2}{2} + \mu\right), \end{aligned} \tag{3.6}$$

where λ are the eigenvalues of the system. It is then necessary that either $A = 0$, or $B = 0$. For the case when $A = 0$, $\mathrm{Re}(\lambda) > 0$ if

$$C_{1u} + C_{2u} < 2, \tag{3.7}$$

and when $B = 0$, after some algebraic simplification, we have that

$$\begin{aligned} &\frac{1}{2}(r_1 + r_2 + 2\lambda)(2\mu + \lambda) = 0, \\ &\implies \lambda_1 = \frac{-r_1 - r_2}{2}, \quad \text{or} \quad \lambda_2 = -2\mu, \\ &\implies \mathrm{Re}(\lambda_{1,2}) < 0 \iff r_i > 0 \quad \text{for} \quad i = 1, 2. \end{aligned}$$

Therefore, this fixed point will be an unstable saddle point, and invadable, if the condition given in (3.7) is satisfied since we have real eigenvalues and one is positive and two are negative; otherwise it becomes a stable uninvadable node since all eigenvalues are real and positive. That is to say that the monomorph will invade the dimorph if the effects of competition on each strain are weak enough.

We now examine equn. (3.5) and calculate the eigenvalues as we have done previously to give the characteristic equation as,

$$\mathrm{CD} = 0$$

where

$$C = -r_u - \lambda,$$
$$D = \Big(r_1\big(1 - C_{u1}\big) - \mu - \lambda\Big)\Big(r_2\big(1 - C_{u2}\big) - \mu - \lambda\Big) - \mu^2,$$

where again it is necessary that either $C = 0$ or $D = 0$. It is clear that when $C = 0$ we have $\mathrm{Re}(\lambda) < 0$ since we have assumed $r_u > 0$. This leaves the only other alternative that $D = 0$, which after making a change of variables, becomes

$$\begin{aligned}&(\tilde{r}_1 - \mu - \lambda)(\tilde{r}_2 - \mu - \lambda) - \mu^2 = 0\\ \iff\ &\lambda^2 + \lambda(2\mu - \tilde{r}_1 - \tilde{r}_2) + \tilde{r}_1\tilde{r}_2 - \mu(\tilde{r}_1 + \tilde{r}_2) = 0,\end{aligned}$$

where $\tilde{r}_1 = r_1(1 - C_{u1})$ and $\tilde{r}_2 = r_2(1 - C_{u2})$. After some algebra we find the remaining eigenvalues are given by

$$\lambda_{1,2} = \frac{\tilde{r}_1 + \tilde{r}_2 - 2\mu \pm \sqrt{(\tilde{r}_1 - \tilde{r}_2)^2 + 4\mu^2}}{2},$$

which then requires a Taylor expansion in μ to approximate the term within the radical — which is reasonable since we assume that $\mu \ll 1$ and that $|\tilde{r}_1 - \tilde{r}_2| >> \mu$. The expansion is truncated at $\mathcal{O}(\mu^2)$ to give an approximation of the eigenvalues as

$$\begin{aligned}\lambda_{1,2} &= \frac{\tilde{r}_1 + \tilde{r}_2 - 2\mu \pm \tilde{r}_1 - \tilde{r}_2}{2},\\ \iff \lambda_1 &= \tilde{r}_1 - \mu, \quad \text{or} \quad \lambda_2 = \tilde{r}_2 - \mu,\\ \iff C_{u1} &< 1 - \frac{\mu}{r_1}, \quad \text{or} \quad C_{u2} < 1 - \frac{\mu}{r_2}.\end{aligned}$$

Since we have assumed that $\mu \ll 1$, the conditions for the dimorph to invade the monomorph are approximately

$$C_{u1} \lessapprox 1, \tag{3.8}$$

and

$$C_{u2} \lessapprox 1. \tag{3.9}$$

Therefore, this equilibrium is also an unstable saddle point because if either conditions (3.8) or (3.9) are satisfied individually, or simultaneously, $C < 0$ and thus one eigenvalue is always negative. Otherwise this will be an uninvadable stable node. Here we should note that there are two entirely separate conditions for the capacity for invasion of the monomorph, each of which corresponds to each strain that comprises the dimorph.

So, we have established that there exist three conditions for invasion, each of which is given by equn. (3.7), giving rise to conditions under which the monomorph can invade the dimorph; and equns. (3.8) and (3.9), which provide conditions under which any dimorphic strains can invade the monomorph. It should not escape our attention that these conditions can be satisfied simultaneously, individually, or not at all. The conditions calculated above are monostable solutions, i.e. when only one stable equilibrium is present. Finding invasion speeds when the system is bistable is a non-trivial problem and is beyond our scope of interest. We shall investigate each of the possible cases in due course, so now turn our attention to calculating the spreading speed of any one of these invasions.

3.2.2 Calculating invasion speeds

Dimorphic invader and weak competition with monomorphic resident

In order to calculate the invasion speed we shall again use the same technique of front propagation into unstable states as we have used previously in chapter 2 [186]. Let us begin by examining the case when the dimorph is posited to invade the territory of the monomorph. Firstly, we linearise equns. (3.3) around the

unstable equilibrium $\mathbf{N}^* = (0, 0, 1)$ to give

$$\begin{aligned}
\frac{\partial v_1}{\partial t} &= D_1 \frac{\partial^2 v_1}{\partial x^2} + \Big(r_1\big(1 - C_{u1}\big) - \mu\Big)v_1 + \mu v_2 + \text{h.o.t.}, \\
\frac{\partial v_2}{\partial t} &= D_2 \frac{\partial^2 v_2}{\partial x^2} + \Big(r_2\big(1 - C_{u2}\big) - \mu\Big)v_2 + \mu v_1 + \text{h.o.t.}, \\
\frac{\partial u}{\partial t} &= D_u \frac{\partial^2 u}{\partial x^2} - r_u C_{1u} v_1 - r_u C_{2u} v_2 - r_u u + \text{h.o.t.},
\end{aligned} \tag{3.10}$$

where again *"h.o.t."* is short-hand for "*higher order terms*". We can then substitute the travelling wave ansatz function for the dimorphic species,

$$v_\iota \propto V_\iota \exp(-i(\omega(k)t + kx)), \quad k, \omega(k) \in \mathbb{C}, \quad \iota = 1, 2$$

where above $i = \sqrt{-1}$, $\omega(k)$ is the dispersion relation, and k is the wave number. A similar substitution is made for u. We should again take heed that, for biologically relevant solutions, k must be purely imaginary — otherwise oscillations close to 0 would produce negative population sizes. Let us now re-parametrise the per capita growth such that $\tilde{r}_i = r_i(1 - C_{ui})$, and make the substitution $\lambda(q) = i\omega(k)$ and $k = iq$ where $q, \lambda(q) \in \mathbb{R}$, which will simplify the algebra and avoid complex numbers. The linearised equations in (3.10) can be written in matrix form as

$$\mathrm{M} = \begin{bmatrix}
\lambda(q) - D_1 q^2 - \tilde{r}_1 + \mu & \mu & 0 \\
\mu & \lambda(q) - D_2 q^2 - \tilde{r}_2 + \mu & 0 \\
-r_u C_{1u} & -r_u C_{2u} & \lambda(q) - D_u q^2 + r_u
\end{bmatrix}, \tag{3.11}$$

which reduces to an eigenvalue problem which must solve the equation

$$\mathrm{M}\mathbf{N} = \mathbf{0},$$

where $\mathbf{N} \in \mathbb{R}^3$ is an eigenvector of the system. The eigenvalues can be found by solving $\det(\mathrm{M}) = 0$, which is equivalent to $\det(\mathrm{M}_{11})\det(\mathrm{M}_{22}) = 0$ where

$$\mathrm{M}_{11} = \begin{bmatrix} \lambda(q) - D_1 q^2 + \tilde{r}_1 + \mu & \mu \\ \mu & \lambda(q) - D_2 q^2 + \tilde{r}_2 + \mu \end{bmatrix}, \qquad (3.12)$$
$$\mathrm{M}_{22} = \begin{bmatrix} \lambda(q) - D_u q^2 + r_u \end{bmatrix},$$

where either $\det(\mathrm{M}_{11}) = 0$ or $\det(\mathrm{M}_{22}) = 0$.

Firstly, we shall check the case when $\det(\mathrm{M}_{22}) = 0$, which when we solve for λ and use the relation given in [186],

$$\left.\frac{d\lambda(q)}{dq}\right|_{q^*} = \frac{\lambda(q^*)}{q^*} \quad \Longrightarrow \; q^2 = -\frac{r_u}{D_u},$$

which is an invalid solution because q will have imaginary parts which violates our requirement for biological consistency. Therefore, the only other option is that $\det(\mathrm{M}_{11}) = 0$. Here we should realise that, once re-parameterised for $\tilde{r}$, this problem has reduced to the same calculation which we have solved previously in Chapter 2. As such, we shall dispense with the calculations and present the possible invasion speeds,

$$\lim_{\mu\to 0} c_{mi} = 2\sqrt{\tilde{r}_i D_i}, \quad i = 1, 2 \quad \text{or} \quad \lim_{\mu\to 0} c_d = \frac{|\tilde{r}_1 D_2 - \tilde{r}_2 D_1|}{\sqrt{(\tilde{r}_1 - \tilde{r}_2)(D_2 - D_1)}}. \qquad (3.13)$$

Similarly we have the same conditions for anomalous invasions as in the chapter 2

$$\frac{\tilde{r}_1}{\tilde{r}_2} + \frac{D_1}{D_2} \geqslant 2 \quad \text{and} \quad \frac{\tilde{r}_2}{\tilde{r}_1} + \frac{D_2}{D_1} \geqslant 2, \qquad (3.14)$$

which hold when $\tilde{r}_i, D_i > 0$. Although the expression for the anomalous speeds remain unchanged, the above conditions need to be considered more carefully in light of competition, and more explicitly, when the competitor causes one strain to have negative net growth.

Dimorphic invader and strong competition with monomorphic resident

The growth parameters in this system are different from our investigation in chapter 2 since $\tilde{r}_i$ can become negative if competition is sufficiently strong enough, i.e. $\delta_{ui} > r_i$; whereas when an occupying resident is not present, r_i is always positive. We used the re-parameterisation $\tilde{r}_i = r_i(1 - C_{ui})$, which can be written more fundamentally as $\tilde{r}_i = r_i - \delta_{ui}$, where δ_{ui} is the mortality of strain i induced by the monomorph u. If we take this into consideration, then we find that q^2 which we used to calculate the invasion speed — see chapter 2 — is given by

$$\lim_{\mu \to 0} q^2 = \frac{\tilde{r}_1 - \tilde{r}_2}{D_2 - D_1},$$

which must be positive. Therefore, without loss of generality, let $\tilde{r}_2 < 0 < \tilde{r}_1$, it is then necessary that $D_2 > D_1$ for $q^2 > 0$ and any invasion speed to be real. Recall from the previous chapter that we determined conditions for the existence of anomalous speeds (Chapter 2.2 equns. (2.23) & (2.24)), which were given by

$$\tilde{r}_1 D_2 + \tilde{r}_2 D_1 - 2\tilde{r}_1 D_1 > 0, \quad \text{and} \quad \tilde{r}_1 D_2 + \tilde{r}_2 D_1 - 2\tilde{r}_2 D_2 > 0, \tag{3.15}$$

or

$$\tilde{r}_1 D_2 + \tilde{r}_2 D_1 - 2\tilde{r}_1 D_1 < 0, \quad \text{and} \quad \tilde{r}_1 D_2 + \tilde{r}_2 D_1 - 2\tilde{r}_2 D_2 < 0, \tag{3.16}$$

which must also be satisfied too. However, if $\tilde{r}_2 < 0 < \tilde{r}_1$, then $\tilde{r}_1 D_2 + \tilde{r}_2 D_1 - 2\tilde{r}_2 D_2 > 0$, is always true and conditions (3.16) can never be satisfied, thus, anomalous invasion speeds will occur provided

$$\frac{D_2}{D_1} > 2 - \frac{\tilde{r}_2}{\tilde{r}_1}. \tag{3.17}$$

Biologically, we will only ever see anomalous invasion speeds with strong competition from the resident if: there is a dispersal-reproduction trade-off be-

tween strains such that dispersal ability is sufficiently different between them; and if the reproductive strain is not competitively excluded, even if the disperser is.

We now have qualitative predictions of an invasion speed when a dimorphic invader supplants a monomorphic resident, let us now progress onto the remaining case when the monomorph is invading the dimorph's habitat.

3.2.3 Calculating invasion speeds: monomorphic invader, polymorphic resident

As we have done previously, we shall linearise equns. (3.3) around the unstable equilibrium $\mathbf{N}^* = (\frac{1}{2}, \frac{1}{2}, 0)$ to give

$$\mathrm{M} = \begin{bmatrix} \lambda(q) - D_1 q^2 - \frac{r_1}{2} - \mu & -\frac{r_2}{2} + \mu & -\frac{r_1 C_{u1}}{2} \\ -\frac{r_1}{2} + \mu & \lambda(q) - D_2 q^2 - \frac{r_2}{2} - \mu & -\frac{r_2 C_{u2}}{2} \\ 0 & 0 & \lambda(q) - D_u q^2 - r_u\Big(1 - \frac{(C_{1u}+C_{2u})}{2}\Big) \end{bmatrix}, \tag{3.18}$$

where again M must solve,

$$\mathrm{M}\mathbf{N} = \mathbf{0},$$

where $\mathbf{N} \in \mathbb{R}^3$ is an eigenvector of the system. The eigenvalues can be found by solving $\det(\mathrm{M}_{11})(\det\mathrm{M}_{22}) = 0$ where

$$\begin{aligned} \mathrm{M}_{11} &= \begin{bmatrix} \lambda(q) - D_1 q^2 - \frac{r_1}{2} - \mu & -\frac{r_2}{2} + \mu \\ -\frac{r_1}{2} + \mu & \lambda(q) - D_2 q^2 - \frac{r_2}{2} - \mu \end{bmatrix}, \\ \mathrm{M}_{22} &= \Big[\lambda(q) - D_u q^2 - r_u\Big(1 - \frac{(C_{1u}+C_{2u})}{2}\Big)\Big]. \end{aligned} \tag{3.19}$$

Finding the associated invasion speed when $\det(\mathrm{M}_{22}) = 0$ is simple, and using the method in [186] we find that the speed is given by

$$\lim_{\mu \to 0} c_{mu} = 2\sqrt{\tilde{r}_u D_u}, \tag{3.20}$$

where $\tilde{r}_u = r_u\big(1 - \frac{C_{u1}+C_{u2}}{2}\big)$. This is the re-parameterised Fisher speed [53, 99] accounting for competition, and is always real when the dimorph can be invaded, i.e. (3.7) is satisfied.

To tackle the remaining case when, $\det(\mathrm{M}_{11}) = 0$, is a non-trivial problem that cannot be solved easily in closed form since the off-diagonal entries are $\mathcal{O}(1)$. If the off diagonals were $\mathcal{O}(\mu)$ we could take the same approach as in chapter 2 by using aymptotic approximation when $\mu \to 0$, although this is not the case and a closed form solution is difficult to attain. However, if $\det(\mathrm{M}_{11}) = 0$, then right eigenvector associated with the system would have a zero entry for u, which would imply that the resident is not present during any invasion, therefore, this case is irrelevant to our investigation. Since it is the instability of the fixed point corresponding to the polymorph being invaded by the monomorph, u will be growing and logically cannot be absent. Therefore, the only remaining possibility is that $\det(\mathrm{M}_{22}) = 0$ and the above speed is the valid case.

3.3 Numerical experiments

Now that we have established analytical predictions for the invasion speeds, it is appropriate to calculate the invasion speeds from numerical experiments to determine if they are in good agreement. The methods used to predict the invasion speed provide a linear approximation that should give a lower bound, but there currently exists no formal proof any species should invade at this speed.

Experiments were conducted in MATLAB 2014a [115] using a finite difference θ-method with $\theta = \frac{1}{2}$ — otherwise known as the Crank-Nicolson scheme [34]. The invasion speeds were calculated by recording 100 time points over the total temporal period and noting the spatial coordinates of the midpoint of the population density invasion front. The speed between each temporal point was calculated — speed $= \frac{\text{distance}}{\text{time}}$ — and an average of those speeds were taken. The results were compared with the analytic predictions and were found to be accurate within the error bounds of the numerical scheme and discretisation steps.

The parameters for simulations of the dimorph out-competing the monomorph are given in table 3.1. Since we are interested in the influence that interspecific competition has on the invasion speed, the interaction matrix for each experiment differ, but take the form

$$C = \begin{bmatrix} 1 & 1 & C_{u1} \\ 1 & 1 & C_{u2} \\ C_{1u} & C_{2u} & 1 \end{bmatrix}, \tag{3.21}$$

where we will note the values as and when we examine each case.

We know from chapter 2 that, with the aforementioned traits, the dimorph will invade anomalously in the absence of a resident competitor. Furthermore, our linear analysis informs us that indeed we could see the same behaviour in this case. Our numerical experiments will inform us whether the non-linear system invades at the same speed predicted by the linearised system. We shall now examine each case of invasion in turn.

3.3.1 Dimorph out-competes the monomorph

We first consider the cases when the dimorph is expected to invade the monomorph.

Both dimorphic strains out-compete the monomorph

In this scenario, our analysis predicts that anomalous invasion speeds should occur when both strains can invade. Parameters have been chosen such that conditions (3.8), and (3.9) are satisfied, and are presented alongside the predicted and calculated invasion speeds in table 3.2

The results can be seen visually in fig. 3.1 panel **a**. Firstly, we notice that both strains invade the monomorph, as predicted by our analysis that gives our conditions for invasion. Secondly, if we compare panel **a** with panel **d**, which is of the reproductive strain in isolation, and panel **e**, which is of the dispersive strain in isolation, we see that in panel **a** the invasion has progressed further then those

Table 3.1: General parameter values used within numerical simulations given in fig. 3.1

Descriptor	**Symbol**	**Values**
Spatial step size	Δx	0.1
Temporal step size	Δt	0.1
Total landscape size	L	[-5000,15000]
Total temporal period	T	5,000
Dispersal coefficients	(D_{v_1}, D_{v_2}, D_u)	$(2, 10, 2)$
Per capita growth rate	(r_{v_1}, r_{v_2}, r_u)	$(0.5, 0.1, 2)$
Mutation rate	μ	0.001
Boundary conditions	∂L	$\frac{\partial \nu(T,L)}{\partial x} = 0, \nu = \{v_i, u\}$
Initial conditions	$v_i = [\min(L), 0]$	$(\frac{1}{2}, \frac{1}{2}, 0)$
	$u = [0, \max(L)]$	$(0, 0, 1)$

Table 3.2: Parameters when both dimorphic strains out-compete the monomorph

Competitive interactions	**Expression**	**Values**
Effect of monomorph on the dimorph	C_{u1}	0.1
	C_{u2}	0.1
Effect of dimorph on the monomorph	C_{1u}	1.1
	C_{2u}	1.1
Analytical invasion speed		
dimorphic speed	$\lim_{\mu \to 0} c_d$	2.5456
strain 1 in isolation	$\lim_{\mu \to 0} c_{m1}$	1.8974
strain 2 in isolation	$\lim_{\mu \to 0} c_{m2}$	1.8974
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	2.5096
strain 1 in isolation	$\bar{c}_{m1}$	1.8542
strain 2 in isolation	$\bar{c}_{m2}$	1.8735

of panels **d** and **e**. Finally, taking the evidence from fig. 3.1 and the calculated speeds in table 3.2, we see that indeed the dimorph invades anomalously when both strains can invade the monomorph.

One dimorphic strain out-competes the monomorph

Next we consider a dimorphic species where one strain is able to invade, but the other is out-competed by our resident monomorphic species. Our analysis indicates that we should only see anomalous invasion speeds when there is strong competition from the monomorph on the dispersive strain. The parameter values used for our simulations of this scenario are given in table 3.4 when it is the dispersive strain under strong competition, and table 3.3 for when the reproductive strain is subject to strong competition.

Both of these cases are illustrated in panels **b**, and **c** in fig. 3.1 respectively. In both panels **b** and **c** we see that the dimorph invades the monomorph, even if one of the strains is out-competed.

Focusing on when the dispersive strain is under strong competition, we can compare panel **b** with the reproductive strain in isolation in panel **d** to find that the invasion has progressed farther. By comparing the entries in table 3.3, we confirm that panel **b** invades anomalously, and that the analytically predicted speeds are in agreement with those calculated from the simulation.

Turning to the remaining case when the reproductive strain is out-competed, we compare panel **c** with the dispersive strain in isolation in panel **e**. Here we find that both invasions proceed at the same rate, which is what our analysis predicted. By checking table 3.4, we can see that the analytically predicted dimorphic speed is undefined, thus the invasion proceeds at the same speed as the dispersive strain in isolation. We should note that the simulations are in good agreement to our analytic calculations.

Table 3.3: Parameter values for when the reproductive strain out-competes the monomorph. Undefined values are given when the invasion speed becomes complex, and thus, are physically meaningless.

Competitive interactions	**Expression**	**Values**
Strong effect on dispersive strain		
effect of monomorph on the dimorph	C_{u1}	1.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	1.1
	C_{2u}	1.1
Analytical invasion speed		
Strong effect on dispersive strain		
dimorphic speed	$\lim_{\mu\to 0} c_d$	2.3262
strain 1 in isolation	$\lim_{\mu\to 0} c_{m1}$	1.8974
strain 2 in isolation	$\lim_{\mu\to 0} c_{m2}$	undefined
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	2.3207
strain 1 in isolation	$\bar{c}_{m2}$	1.8542

Table 3.4: Parameter values for when the dispersive strain out-competes the monomorph. Undefined values are given when the invasion speed becomes complex, and thus, are physically meaningless.

Competitive interactions	**Expression**	**Values**
Strong effect on reproductive strain		
effect of monomorph on the dimorph	C_{u1}	1.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	1.1
	C_{2u}	1.1
Analytical invasion speed		
Strong effect on reproductive strain		
dimorphic speed	$\lim_{\mu\to 0} c_d$	undefined
strain 1 in isolation	$\lim_{\mu\to 0} c_{m1}$	undefined
strain 2 in isolation	$\lim_{\mu\to 0} c_{m2}$	1.8974
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	1.8641
strain 2 in isolation	$\bar{c}_{m2}$	1.8735

3.3.2 Monomorph out-competes the dimorph

Now we shall explore when a monomorph invades a dimorphic resident's habitat. The parameters for this experiment are presented in table 3.5. Note that for these experiments we change the initial conditions such that $L = [-13000, 13000]$, all other parameters remain unchanged.

The results of this simulation are illustrated in fig. 3.2 panel **e**. Our experiment has shown that the monomorph invades the dimorph and displaces it as our analysis predicted. Furthermore, the invasion travels at the analytic speed given in (3.20) and shown in table 3.5.

3.3.3 Both dimorph and monomorph coexist

Our mathematical analysis and numerical experiments have so far shown us that the conditions for either species to be invaded are determined by the strength of competition applied from the other, e.g. the dimorph will invade provided that competition exerted by the monomorph is weak enough. It is therefore possible that if the strength of competition each species experiences is low enough, both can invade each other's habitat.

We shall note that finding any closed form expression for the coexistence equilibrium and its stability is non-trivial. It is for this reason that we shall forego any mathematical analysis of their existence and stability. Nevertheless, it is only the unstable states which provide invasion speeds, which we have calculated and are testable. Therefore, we shall consider these cases using simulations with the knowledge that each species can be invaded by the other.

Dimorph faces weak competition

When competition between both species is weak enough, we find that all of the conditions for the dimorph to invade given in section 3.2.1 are satisfied, which informs us that a counter-invasion is possible. For this scenario, the parameters given in tables 3.1 and 3.8, provide a coexistence solution which are found

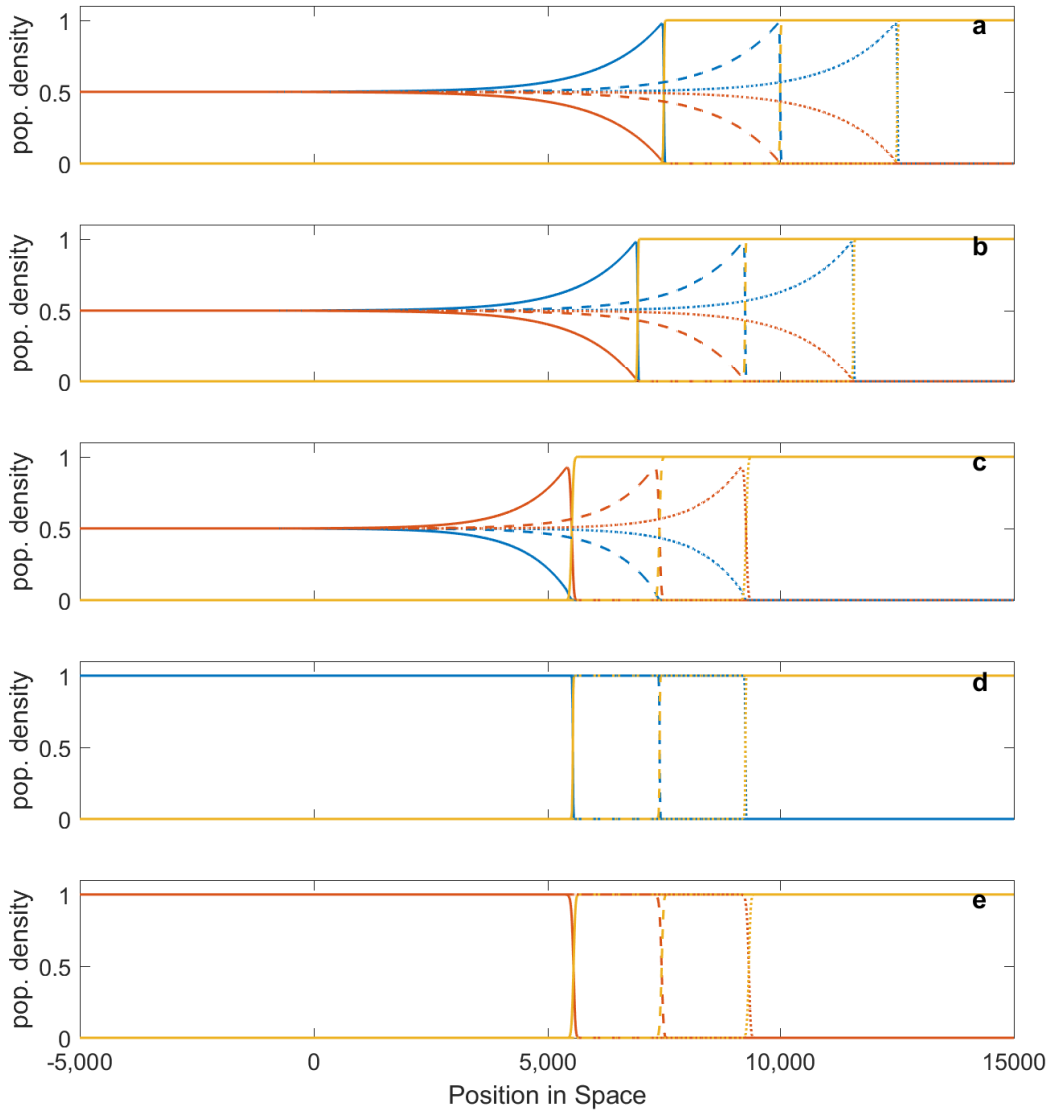

Figure 3.1: Dimorphic invasion of habitat occupied by a monomorphic resident which is then displaced. Population density is given on the y-axis. Position in space is given on the x-axis. Parameters for reproductive strain were $D = 2$, $r = 0.5$ (blue) , dispersive strain $D = 10$, $r = 0.1$ (orange), both initial conditions are $(\frac{1}{2}, \frac{1}{2}, 0) = [-5000, 0]$. Parameters for the monomorph are $D = 2$, $r = 2$ (yellow) with initial conditions $(0, 0, 1) = [0, 15000]$. Invasion at 3 equidistant time points are shown in chronological order and depicted at $t = 3,000$ (solid), $t = 4,000$ (dashed), and $t = 5,000$ (dotted) lines respectively. Panel a shows the case when both strains can invade and the monomorph cannot. Panel b shows the case when only the reproductive strain can invade. Panel c shows the case when only the dispersive strain can invade. Panel d shows the reproductive strain in isolation. Panel e shows the dispersive strain in isolation. Mutation parameters were $\mu_{12} = \mu_{21} = 0.001$.

numerically. The results of this simulation are illustrated in fig. 3.2 panel **a**.

We see from panel **a** that both species freely invade the other and coexist. To address the dimorphic invasion, since we used the same values as in section 3.3.1 when the monomorph could not invade, the invasion speeds can be compared by examining tables 3.6 and 3.2. We find that the calculated invasion speeds differ within an order of the numerical scheme which suggests they are the same and anomalous.

Attending to the monomorphic invasion in the same way, comparing tables 3.6 and 3.5 show a similar result that invasion speeds are the same for both monomorphic invasions.

Therefore, both species invade at the linearly predicted speeds given by equn (3.20) for the monomorph, and equn (3.13) irrespective of whether the other species is able to invade it's territory or not.

One dimorphic strain faces strong competition

Consider the scenario when a monomorph is invading a dimorph, but only one strain is capable of counter-invading. The parameters used within the experiments are given in table 3.7 when the dispersive strain is facing strong competition, and in table 3.8 when it is the reproducer experiencing stiff competition.

Results of these simulations can bee seen in panels **b** and **c** of fig. 3.2. The first feature we notice is that both species invade into the other's territory, but the out-competed strain exists in the population at low densities comparable to the mutation rate, i.e. $v_{out-competed} = \mathcal{O}(\mu)$. Secondly, the dimorphic invasion when the dispersive strain is out-competed, panel **b**, is invading faster than when the reproductive strain is out-competed, panel **c**. Finally, upon consultation with tables 3.7 and 3.8, we see that both cases invade at to their respective linearly predicted invasion speeds, as does the monomorph.

Table 3.5: Parameter values for when the dimorph is out-competed by the monomorph

Competitive interactions	**Expression**	**Values**
Effect of monomorph on the dimorph	C_{u1}	1.1
	C_{u2}	1.1
Effect of dimorph on the monomorph	C_{1u}	0.1
	C_{2u}	0.1
Analytical invasion speed		
monomorphic speed	c_{mu}	3.7947
Average speed from simulations		
monomorphic speed	$\bar{c}_m$	3.5119

Table 3.6: Parameter values for when both monomorph and dimorph can invade equally

Competitive interactions	**Expression**	**Values**
effect of monomorph on the dimorph	C_{u1}	0.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	0.1
	C_{2u}	0.1
Analytical invasion speed		
dimorphic speed	$\lim_{\mu \to 0} c_d$	2.5456
monomorphic speed	c_{mu}	3.7947
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	2.5082
monomorphic speed	$\bar{c}_{mu}$	3.5119

Table 3.7: Parameter values for when the dispersive strain is out-competed, but the monomorph and reproducer can mutually invade

Competitive interactions	**Expression**	**Values**
Strong effect on dispersive strain		
effect of monomorph on the dimorph	C_{u1}	0.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	1.1
	C_{2u}	0.1
Analytical invasion speed		
dimorphic speed	$\lim_{\mu\to 0} c_{m2}$	2.3562
monomorphic speed	c_{mu}	3.7947
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	2.3194
monomorphic speed	$\bar{c}_{mu}$	3.5119

Table 3.8: Parameter values for when the reproductive strain is out-competed, but the monomorph and disperser can mutually invade

Competitive interactions	**Expression**	**Values**
Strong effect on reproductive strain		
effect of monomorph on the dimorph	C_{u1}	1.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	0.1
	C_{2u}	0.1
Analytical invasion speed		
dimorphic speed	$\lim_{\mu\to 0} c_d$	1.8974
monomorphic speed	c_{mu}	3.7947
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	1.8510
monomorphic speed	$\bar{c}_{mu}$	3.5119

Founder control

This then leaves us with the remaining case of when competition experienced by both species is strong enough such that neither can invade the other. We should note here that the linear analysis to calculate an invasion speed is not applicable for this scenario, therefore we shall explore it using simulations only. Parameters used in this experiment are given in table 3.9 and results in fig. 3.2 panel **d**. We see from panel **d** that there is very little movement, and from table 3.9 a very slow invasion by the dimorph. In fact, with a different choice of parameters we would see a monomorphic invasion, but again very slowly — numerical experiment omitted. Here we highlight that since the equations are non-linear, and our analysis uses a linear approximation, we are missing some of the behaviour exhibited by the equations. In this case, it is non-linearities from behind the wavefront that are propagating the solution forward — a pushed front, rather than a pulled front like the previous cases. Therefore, we cannot make predictions using the methods discussed in this chapter.

These numerical experiments have shown us that in all of the examples shown the invasion travelled at a speed predicted by our linear analysis. Our numerical experiments have confirmed that anomalous invasion speeds are still possible, even in the presence of a competitor species. Furthermore, anomalous invasion speeds persist only when the reproductive strain is able to invade, regardless of whether the dispersive strain can too. Our simulations have also shown that the non-linear system follow the linearly predicted invasion speeds laid out in our analysis, which is true even when the invader can itself be invaded. Overall, the numerical experiments have validated that the linear analysis is a good approximation for the invasion speed of the non-linear system.

3.4 Discussion

Our investigation has uncovered aspects of polymorphic invasions of habitat occupied by a resident species. We have derived conditions and invasion speeds when each species could invade the other. We have established that anomalous invasion speeds for a dimorphic species are still possible, even in the presence of a resident. Furthermore, we also found that, after we re-scaled the growth parameters, the relations for the invasion speeds, and the conditions upon them being realised, were the same as when no resident is present. This naturally brought us to cases when one of the dimorphic strains could invade and the other could not, where we have seen that anomalous invasions are still possible. Surprisingly, this was also true when the dispersive strain has negative net growth. Finally, we explored situations when both species were able to invade each other.

Previous theoretical research has shown that invasions are slower when a resident is present. The effect of a resident competitor has been shown to slow the spread of an invader in previous models using reaction-diffusion equations [129, 109]; integro-difference equations [77, 108, 193]; and in cellular automata models [19]. This trend has been reflected in empirical systems too, whereby an invasion of an exotic annual grass has been slowed by competition from native perennial species of grasses in California [158]; and on the Big Island of Hawai'i where the spread of cane tibouchina (*Tibouchina herbacea*) was shown to be slower when competing against forest communities [4]. In our model, mortality induced by competition from the other species δ_{ij} reduces the net growth rate, therefore slowing the speed of invasion.

Our first significant outcome is that anomalous invasions are robust in the presence of a resident competitor. In chapter 2 we found that when a polymorphic species invaded vacant habitat anomalous invasions could occur if the correct conditions were met. Here, under the re-parameterisation $\tilde{r}_i = r_i(1 - C_{ui})$, we find an identical set of invasion speeds and conditions given in (3.13), and (3.14). Therefore, all of our conclusions from the previous chapter will also hold in this case.

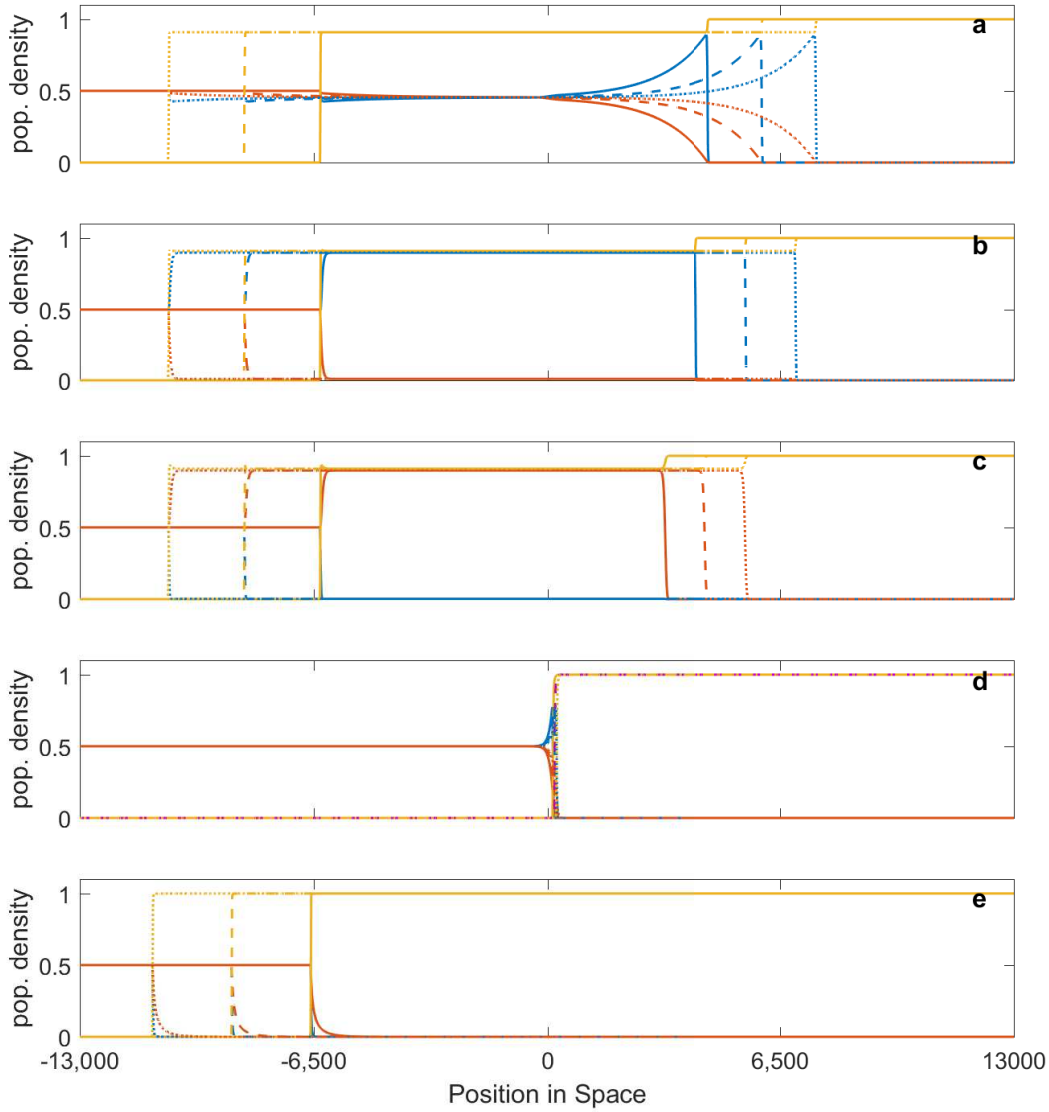

Figure 3.2: Monomorphic invasion of habitat occupied by a dimorph. Population density is given on the y-axis. Position in space is given on the x-axis. Parameters for reproductive strain were $D = 2$, $r = 0.5$ (blue) , dispersive strain $D = 10$, $r = 0.1$ (orange), both initial conditions were $(\frac{1}{2}, \frac{1}{2}, 0) = [-13000, 0]$. Parameters for the monomorph are $D = 2$, $r = 2$ (yellow) with initial conditions $(0, 0, 1) = [0, 13000]$. Invasion at 3 equidistant time points are shown in chronological order and depicted at $t = 1,800$ (solid), $t = 2,400$ (dashed), and $t = 3,000$ (dotted) lines respectively. Panel a shows the case when the dimorph can counter-invade the monomorph. Panel b shows the case when the reproductive strain can counter-invade the monomorph. Panel c shows the case when the dispersive strain can counter-invade the monomorph. Panel d shows the case when neither monomorph nor dimorph can invade the other. Panel e shows the case when a monomorph invades displacing the dimorph. Mutation parameters were $\mu_{12} = \mu_{21} = 0.001$

When no competitor is present all strains can invade, but this is not always the case when competition from the resident is strong enough. However, we found that despite the resident competitively excluding one strain, the dimorph could still invade provided the other strain was not out-competed. Furthermore, we also revealed that anomalous invasion speeds persisted when the dispersive strain was out-competed, that is, had negative net growth. This was not true of the opposite case when it was the reproductive strain that was out-competed, suggesting that reproduction is more critical than dispersal for anomalous invasions when the species is in competition.

We can use this to highlight an important consideration for invasion management. Suppose a dimorphic species is invading where one strain is seemingly competitively excluded, a strategy may be adopted erroneously based solely on the successful strain since it appears to be the only survivor, where in reality an anomalous invasion may be occurring. The results of our model suggest that this could be a possibility, and we can speculate on its generality and limitations.

Our model reflects the case when there are two species, one dimorphic and the other monomorphic. However, the structure of the system will easily generalise to the case when there are N-species, each with a general degree of polymorphism. Although the speeds provided in chapter 2 have been proven to represent the non-linear system by Girardin [62], the results cannot simply be extended without proof. Therefore, we can only suggest that these results are true after witnessing strong agreement with our numerical simulations. Our results are valid for haploid asexual individuals so we should note that further investigation is required to find if these results can extend to a model that accounts for sexual recombination of diploid organisms – which we shall address in chapter 5. It should also be noted that our values of r_i and D_i have been arbitrarily chosen to illustrate the dispersal-reproduction trade-off required for anomalous invasion. Therefore, we would like to find how general these types of trade-off are in nature, which we shall be investigating in the proceeding chapter.

Table 3.9: Parameter values when either species cannot invade the other

Competitive interactions	**Expression**	**Values**
effect of monomorph on the dimorph	C_{u1}	0.1
	C_{u2}	0.1
effect of dimorph on the monomorph	C_{1u}	0.1
	C_{2u}	0.1
Analytical invasion speed		
dimorphic speed	$\lim_{\mu \to 0} c_d$	undefined
monomorphic speed	c_{mu}	undefined
Average speed from simulations		
dimorphic speed	$\bar{c}_d$	0.0877

Chapter 4

Empirical evidence for dispersal trade-offs: a meta-analysis

4.1 Introduction

A critical element of this thesis is the role of dispersal life history trade-offs within range expansions, for which quantitative estimates are lacking. In life history theory, dispersal is considered to be an expensive trait that, to be maintained, must do so at the expense of other traits [12, 155]. Trade-offs within organisms between dispersal and other traits have been proposed with strong theoretical backing. However, what is noticeably absent is collective empirical evidence for dispersal trade-offs. Much of the literature within this topic is theoretical in nature, and the majority of the review articles have small sections that cite a small number of studies [78, 22, 102]. Empirical investigations that explicitly examine dispersal trade-offs are rare and within very few taxonomic groups – primarily crickets (*Orthoptera*) [66, 203, 123] and butterflies (*Lepidoptera*) [86, 75, 81] – which makes any evidence for the existence of dispersal trade-offs tenuous.

The notable absence of empirical evidence is unsurprising. It is notoriously difficult to accurately quantify an organism's dispersal ability outside of the few

recognised proxies that are currently used. Furthermore, any such proxy is open to statistical confounding from other fitness traits, e.g. body size or condition is used as a proxy for dispersal, but is also used as a proxy for competitive effects [33, 113]. As mentioned previously, few studies explicitly examine dispersal trade-offs, so any potential evidence is spread throughout the literature within studies that are outside the field of dispersal, and by extension, life history trade-offs associated with dispersal. This absence of evidence poses a problem for the development of theories that examine dispersal trade-offs. Without tangible evidence the theory cannot be validated or invalidated, stunting the advancement of life history theory.

A solution to the problem of forming quantitative effects from information scattered throughout the literature is to use meta-analysis. A notable problem with narrative reviews are that they do not quantify the collective effects of the subject being reviewed. A review will also rely on the author's knowledge of the subject, which may have gaps and be biased towards certain areas of literature. The process of a meta-analysis looks to address these issues by systematic review and statistically determining an overall effect size across all studies included. The technique of meta-analysis has been used extensively within clinical medicine to aggregate results from many different studies to determine the collective effect of a treatment with great success [46]. There is a distinguished body of literature in ecology that has used meta-analysis. Meta-analytic techniques have been used to determine relationships between adaptive and neutral genetic and ecological structure and functioning; determining the roles of natural selection and genetic drift within variation of immune genes; and predicting the metabolic rate of seabirds [197, 172, 44]. In a similar study to this, meta-analysis has also been used to determine trade-offs between flight capability and reproduction in cricket species, which implies this type of analysis is well suited to our needs [66].

We aim to address the lack of quantitative dispersal trade-offs by conducting a systematic review of the literature for available data and performing a meta-analysis to determine the prevalence of general dispersal trade-offs. In doing so,

we aim to build a database of the aggregated data on dispersal trade-offs to address the issue of sparsely populated data that is a substantial barrier to research on this topic. We also aim to determine the existence of dispersal trade-offs from the data within this database to determine if current theories are correct to assume their existence. In our previous chapters 2 & 3 the shape of the dispersal-reproduction trade-off was critical for the existence of anomalous invasion speeds, therefore, a key outcome for our theory would be to use this meta-analysis to determine if the dispersal trade-off conjecture is valid.

We surveyed the literature and extracted the available data using a robust inclusion criteria. The data was then categorised into factors we proposed would have an effect on any dispersal trade-off: dispersal type, fitness proxy, sex, species class, climate classification, and environmental lineage. We analysed the data using meta-analytic techniques to test whether there was an overall effect of a dispersal trade-off, and then tested if any of the aforementioned factors could explain it. Since the data within the explanatory factors was clustered, any of the factors that expressed significant explanatory evidence for a trade-off were then tested again while controlling for the remaining factors.

We show that there is empirical evidence that can be examined for dispersal trade-offs from the construction of the datasets, and we found from our analysis that evidence for a trade-off exists. Our study adds to the evidence for a vital component of dispersal life-history theory that has had few empirical contributions, furthermore, giving weight to the theoretical conjecture that dispersal is subject to a life history trade-off.

4.2 Data collection

Our review will extend to cover both natural and agricultural populations, or those subject to human selection pressures. We also include all study types, such as long-term multi-generational populations and selection lines. We have included as many study types and organisms as we could account for because we would like

to draw general conclusions about the existence of dispersal trade-offs. By having a broad range of study types we can determine if the experimental procedure has an impact on any dispersal trade-off. In order to investigate trade-offs, all relevant studies included both measures of dispersal and other quantitative fitness traits which we categorise in section 4.2.4.

4.2.1 Search terms

We developed a comprehensive list of search terms which were used to search the database ISI Web of Science Core Collection for relevant literature. We opted to restrict the search to this database alone to reduce potential search bias from publishers' own databases (e.g. Scopus, Science Direct). We opted to exclude Google Scholar since the algorithm takes previous searches into account and could bias search results.

We are investigating the prevalence of trade-offs between dispersal and other non-dispersal life-history traits. Therefore, we search the database by separating each component of our investigation with associated terminology and synonyms familiar within those fields.

The search operations within ISI Web of Science follow Boolean rules for combining terms. For our search, the following operations were required

- "X" — Exact phrase. The search will find the exact phrase X within quotations.

- X* — Wildcard. The search will find the phrase X and will also include any characters after the asterisk, e.g. disp* will return dispersal, dispersers, etc.

- NEAR\X — Proximity operator. The search will find phrases linked by this operator within a proximity of X words, where X is an integer.

The search we conducted used the terms

1. Dispersal terms
   dispers* OR colonis* OR coloniz*

2. Trade-off terms
   (fitness AND cost*) OR "trade-off" OR "trade-offs" OR ((trait* OR life-histor*) NEAR\10 (correlat* OR relationship*)) OR (phenot* NEAR\10 (correlat* OR relationship*))

3. Stand-alone terms

   (a) "dispersal polymorphism"

   (b) "dispersal morph"

   (c) (dispers* OR colonis* OR coloniz* NEAR/5 cost*)

We then used Boolean operations on each of the searches within the database for possible articles for inclusion and recorded the results in table 4.1. The search was conducted on 24/MAR/2017. The final total for articles that were assessed under our criteria was 5,392.

Table 4.1: Summary of the number of articles returned from the database search.

Search terms	**Number of returned articles**
1.	781,533
2.	146,347
3. (a)	45
3. (b)	5
3. (c)	1,322
1. AND 2.	4,219
(1. AND 2.) OR 3. ((a) OR (b) OR (c))	5,392.

4.2.2 Study inclusion criteria

Relevant subjects

We developed a robust inclusion criteria to assess articles for inclusion in our analysis. We allowed for individuals or populations to be either prokaryote or eukaryote, but excluded humans and viral-type organisms. Due to the nature of viral life we feel that their life histories cannot be interpreted in the same manner as other organisms. We exclude humans due to their complex dispersal behaviours that cannot be easily quantified within this analysis. We exclude organisms investigated in medical or clinical studies focusing on treatment of ailments because dispersal patterns are not interpretable in a way that is consistent with natural pattens of pathogen dispersal in other host systems.

To identify trade-offs, subjects had to be within the same species and measurements taken comparing individuals (when both traits were measured within the same individual), or comparing populations (when both traits were not necessarily measured within the same individuals). Since we are interested in intraspecific variation of traits, comparative multi-species studies were excluded during this study.

Relevant trait measurements

Here we define the measurements of traits within studies that were relevant for inclusion. Studies had to quantify dispersal alongside other non-dispersal fitness traits. We allowed for the quantification of traits to be either continuous or discrete. To select relevant trait measurements we had to define a scale of quality. Dispersal measurements were categorised into three types:

- *Realised dispersal* – When dispersal has been physically quantified, i.e. genetic markers have established dispersal, or distance has physically been measured.

- *Validated dispersal proxy* – A proxy for dispersal which has been validated to

directly influence dispersal in previous studies where evidence of this exists. That is, there exists a previous study which has associated the proxy trait to dispersal.

- *Non-validated dispersal proxy* – A proxy for dispersal which has been theoretically proposed to influence dispersal, but evidence for this is lacking. That is, there were no other previous studies which link the proxy trait to dispersal.

This categorisation also pertained to the measurement of fitness traits such that:

- *Validated fitness proxy* – A trait which has been validated to directly affect fitness in previous studies where evidence of this exists.

- *Non-validated fitness proxy* – A trait which has been theoretically proposed to have an impact on fitness, but evidence for this is lacking. That is, there were no other studies which linked the proxy to fitness that we were aware of.

Since there are no realised measures of fitness, we have made the distinction between validated and non-validated as a validated trait when there exists more than 1 previous study linking that trait to fitness, otherwise the trait was considered non-validated. The proxies of offspring or egg counts were declared validated without a reference since they are "common" proxies for fecundity throughout the literature. Longevity was also treated in this way, but as a fitness category in its own right. However, longevity could only be counted as natural sensence, since any organism suffering mortality from an external factor would not convey a trade-off in a comparable way to the other factors under investigation.

Due to the nature of our scientific question, we had to define traits that we believed should be excluded. We excluded traits associated with dispersal which had confounding factors with other fitness traits, such as body size. Ad-

ditionally, a special case of a botanical trade-off exclusively between seed size and seed number, were excluded. This is occasionally presented as a case of a dispersal trade-off, and has been proposed theoretically, but the empirical evidence is lacking. As such, only when a clear validated proxy existed, such as wind-dispersal, were seed mass and seed number included. We exclude any probabilities of dispersal (dispersal propensity) for organisms since our focus is on direct empirical evidence, rather than inferences. We also exclude cases when only data relating to traits associated with temporal dispersal (e.g. seed banks, delayed dispersal, seed gut passage time) since the physical dispersal trait has not been quantified. We do not include any study which examines migration. We define migration as the process of annual movement by a population from one location to another. We consider this to be different from the dispersal process because migration involves the periodic movement of the entire population, whereas dispersal involves movement relative to that population's natal location. We only included survival/mortality as a non-dispersal fitness trait in the case of senescence (longevity), since any other measurement of mortality could be confounded by non-life history events, e.g. predation.

4.2.3 Assessment of study relevance

We assessed each study for inclusion within the meta-analysis using a hierarchical approach. We assessed articles for relevancy by first examining the article titles; then by abstracts; and finally, those that remained were examined using the article's full text for relevant data to be extracted. When assessing the articles they were either accepted to the next level of assessment, or were rejected and no longer considered for further scrutiny.

The inclusion criteria remains valid for any other researcher wishing to repeat our analysis and is detailed below

Article title assessment

Articles are assessed for inclusion within the study according to the following:

- Articles whose titles indicate the subject of the study is a non-human organism should be retained.

- Articles whose titles indicate a medical subject area focussing on humans will be excluded.

- Articles whose titles indicate a clearly non-biological subject area (e.g. engineering) will be excluded.

Abstract assessment

Articles are then assessed for inclusion based on the following rules and in chronological order:

1. Are the dispersal traits of individuals or populations (mean phenotypes) quantified?

   (a) Traits associated with dispersal must be measured.

   (b) Any article which explicitly measures a dispersal trade-off should be retained.

   (c) Any article which has traits described in the abstract as dispersal traits or dispersal phenotypes (including binary "dispersers") should be retained.

   (d) Any article which has traits that affect the mobility/distribution of the organism should be retained.

   (e) Any article pertaining to a cost for dispersal should be retained.

   (f) Any article pertaining to migration should be excluded. However, if

the study investigates dispersal within a migratory species it should be retained for full text analysis.

2. Articles which measure the dispersal traits at the individual and population level within the same species are to be retained.

   (a) Any article which does not make an explicit or implicit distinction should be retained.

   (b) Any article which has multiple species where only one species is the main focus are to be retained, e.g. a parasite's dispersal ability is measured where there are multiple hosts.

3. Studies will only be included if they represent primary empirical research, with the following exceptions

   (a) Reviews (secondary research) will be retained if they are relevant to the subject area of this meta-analysis, and examined for additional primary citations.

   (b) Any theoretical/statistical or modelling articles which explicitly state empirical data has been fit, or otherwise use empirical data should be retained.

Any article for which any of the above criteria were uncertain were retained to the next level of assessment. We accepted the uncertain articles because false positives would only become apparent at the full text analysis stage.

Full text assessment

Articles are then subjected to a full text examination

1. Are other non-dispersal related fitness traits also measured?

(a) Non-dispersal related traits must be objective and quantified alongside dispersal traits.

i. Any article which explicitly measures dispersal trade-offs should be retained.

ii. Any article which only examines dispersal traits should be dismissed.

2. Articles will only be included if they have appropriate data for calculation in the meta-analysis.

(a) Traits can be either continuously distributed, or have discrete classifications — e.g. winged and wingless organism. For our analysis we require data presented as means with standard deviations or standard errors, or Pearson's correlation coefficient is presented, or when either can be calculated from the data.

(b) Studies that examine individuals must have $n \geq 10$ subjects within a study.

(c) Studies that examine populations must have $n \geq 3$ populations within a study.

(d) Articles that present both raw data and/or summary statistics are to be retained.

(e) Any trait which has been estimated shall be excluded, such as traits given as probabilities, or any trait which has been given as an output from a predictive model.

(f) All article authors should be contacted in an attempt to attain raw

data used within each study.

Review articles

When an appropriate review was encountered, the articles referenced within were then assessed by the above criteria for abstracts and full text. We excluded the title, first, and second levels of article assessments since the review will have already passed those criteria. This meant that we assessed whether the articles were from primary empirical research (lvl. 3). If another review was collected in this manner it was not subjected to further assessment and was excluded to stop a circular process. Of the articles in our original search, 28 were reviews from which we collected all of the available citations for assessment. From the review articles we gained another 2,557 articles to assess after duplicates from the original search were removed.

Resulting assessment

We applied the above assessment to the articles found in the search and documented the number at each level of scrutiny. Of the 7,949 articles assessed, the remaining number of 224 were retained for data extraction. The assessment of the articles from the original search was solely conducted by Vincent Keenan (VAK) for all of the articles. To ensure that the inclusion criteria was robust and objective, a random sample of 500 articles – approximately 9% of the total number – were assessed by Amy Eacock (AE) using the same inclusion criteria. We found that both VAK and AE accepted and rejected a similar amount of articles which was quantified using the Kappa (κ) coefficient of agreement applied to independent assessments of article subsets [30].

The assessments of review articles was evenly split between VAK (1,279 articles) and AE (1,278). A κ - coefficient was not calculated for the review assessments since acceptable concordance of the inclusion criteria had been established with those conducted from the initial search. If a review article did not have an available abstract it was excluded. The results of the entire assessment procedure are summarised in table 4.2.

Table 4.2: Articles remaining and κ – coefficients at each level of the inclusion criteria assessment

Original search			**Review citations**
Assessment level	**Number of articles retained**	**κ coefficient from rand. samp. of 500**	**Number of articles retained**
Article title	3,721	0.882	
Article abstract lvl. 1	1,219	0.428	
Article abstract lvl. 2	821	0.807	
Article abstract lvl. 3	498	0.631	570
Full text	171	N/A	53
Total articles for extraction assessment	**224**		

4.2.4 Data extraction

After we filtered the articles the data available within was then extracted. We attempted to attain as much of the raw data used within each article to calculate our own summary statistics. We contacted all of the article authors in an attempt to acquire as much of the raw data as was possible, unless it was freely available within an online repository. If we could not access the raw data, then we retained the reported summary statistics, or we exported data from reported graphical outputs using "PlotDigitiser" (GNU Public Licence version 2.0).

Data types

For each article examination which passed the above inclusion criteria, we had to categorise the various trait measurements into larger aggregated categories. When extracting the available trait data we categorised them into four groups of data types: dispersal and fitness both continuous, dispersal and fitness both discrete, dispersal continuous and fitness discrete, and dispersal discrete and fitness continuous. For each article, we collected a suite of meta-data that were identified as potential sources of heterogeneity within the meta-analysis. Since one of the main aims of this project was to publish our collected data within an online repository for other researchers to access. Meta-data that is potentially useful for others was collected, even if it was not used within our analysis.

From our extraction process we were only able to attain 3 data types: continuous dispersal and non-dispersal data, discrete dispersal and continuous non-dispersal data, and discrete dispersal and non-dispersal data. The data set that consisted of both discrete trait measures of dispersal and non-dispersal traits had only 2 studies within it, which is too few to perform any analysis and was not examined any further. We therefore had two data sets that could be subject to meta-analysis: one with dispersal and non-dispersal traits as continuous measurements; and one with dispersal traits as discrete, and non-dispersal traits as continuous measurements.

Effect size classification

In order for the data to be analysed, an effect size had to be calculated to make studies comparable. For the discrete dispersal trait and continuous non-dispersal trait data – hereafter known as the binary data set – we extracted means, standard deviations or standard errors, and sample sizes of each study. When dispersal was discrete the data was binary, that is, a disperser and non-disperser – or lesser disperser – was identified categorically and means standard deviations were either taken or calculated from the continuous non-dispersal trait data. We then took a standardised mean difference between dispersers and non-dispersers as an effect size, details of which will be explained later. For both continuous dispersal and non-dispersal trait data – hereafter known as the continuous dataset – a Pearson's correlation coefficient was either extracted or calculated between dispersal and non-dispersal traits as our effect size.

The final number of data points for the meta-analysis, that is studies from all articles, were 36 for the continuous data set and 114 for the binary dataset.

There exist methods to convert one effect size into another, such as a standardised mean difference into a correlative, but would be ill-advised in our case. This can be explained by the fact that an observational study reporting a correlation is indeed different from a study reporting a mean effect. The difference can be summarised as a correlation which measures the change in one trait relative to the other, for example dispersal and reproduction, whereas a standardised mean difference measures difference between two groups, for example: dispersers and non-dispersers reproductive ability. We decided to err on the side of caution and kept both data sets separated during analyses.

Categorisation of traits

For the purpose of our analysis we grouped traits into higher level categorisations. We examined the dispersal traits from our data set and categorised dispersal into two distinct types:

- Physical investment – dispersal traits which require a physical morphological investment.

- Facultative – dispersal traits that are related to behavioural decisions, or distance travelled.

We justify this decision by reasoning that although all organisms have to make some form of facultative decision to disperse, not all organisms invest heavily in dispersal apparatus, such as wings.

For non-dispersal fitness traits we assessed the reported unit of measurement and categorised them into three categories:

- Fecundity – any trait which is directly attributable to procreation.

- Longevity – the length of time taken until natural death.

- Competition – any trait which would affect how that organism performed during its lifetime.

We appreciate that there is an argument that fecundity and competition could be considered indistinguishable with regards to fitness, that is, both traits encompass the contribution of that organism to future generations. However, we justify this distinction by the type of data that was collected.

Climate data from where each organism originated from was collected and aggregated into a higher climate classification. We used the Köppen-Geiger (KG) climate classification system for identifying climatic conditions [133] and categorised them into five categories

- A – Tropical.

- B – Dry.
- C – Temperate.
- D – Continental.
- E – Polar.

We also collected data of the environmental lineage of the study subjects. These were categorised into three categories

- Common garden – Data collected was from experiments from organisms that were taken from a common garden environment, and were born within these conditions.
- Laboratory stock – Data collected was from experiments from organisms that were taken from a laboratory population, and were born within these conditions.
- Source environment – Data collected was from experiments conducted on organisms that were from, or immediately taken from, their natural environment.

We chose not to make the distinction of how many generations from the common garden or lab stocks had been reared in these conditions because this information was not always reported in the articles.

Potential sources of heterogeneity

When conducting meta-analysis it is important to identify any potential sources of heterogeneity within effect sizes. Our sources of heterogeneity that are our explanatory variables were chosen *a priori* for our analysis so as not to influence our hypothesis testing and are given by

- Dispersal type – whether the dispersal trait was a physical investment or facultative.

- Fitness type – whether the non-dispersal trait was fecundity, longevity, or competitive.

- Species class – the taxonomic species class.

- Sex – the sex of the study subjects.

- KG climate class – which climate classification the subjects were, or were descended from.

- Environmental lineage – whether the study subjects were from common garden, laboratory stock, or source environment.

After extracting the data from the articles, we were able to calculate proportions of our explanatory variables for sources of heterogeneity within the continuous data set in figure 4.1, and within the binary data set in figure 4.2.

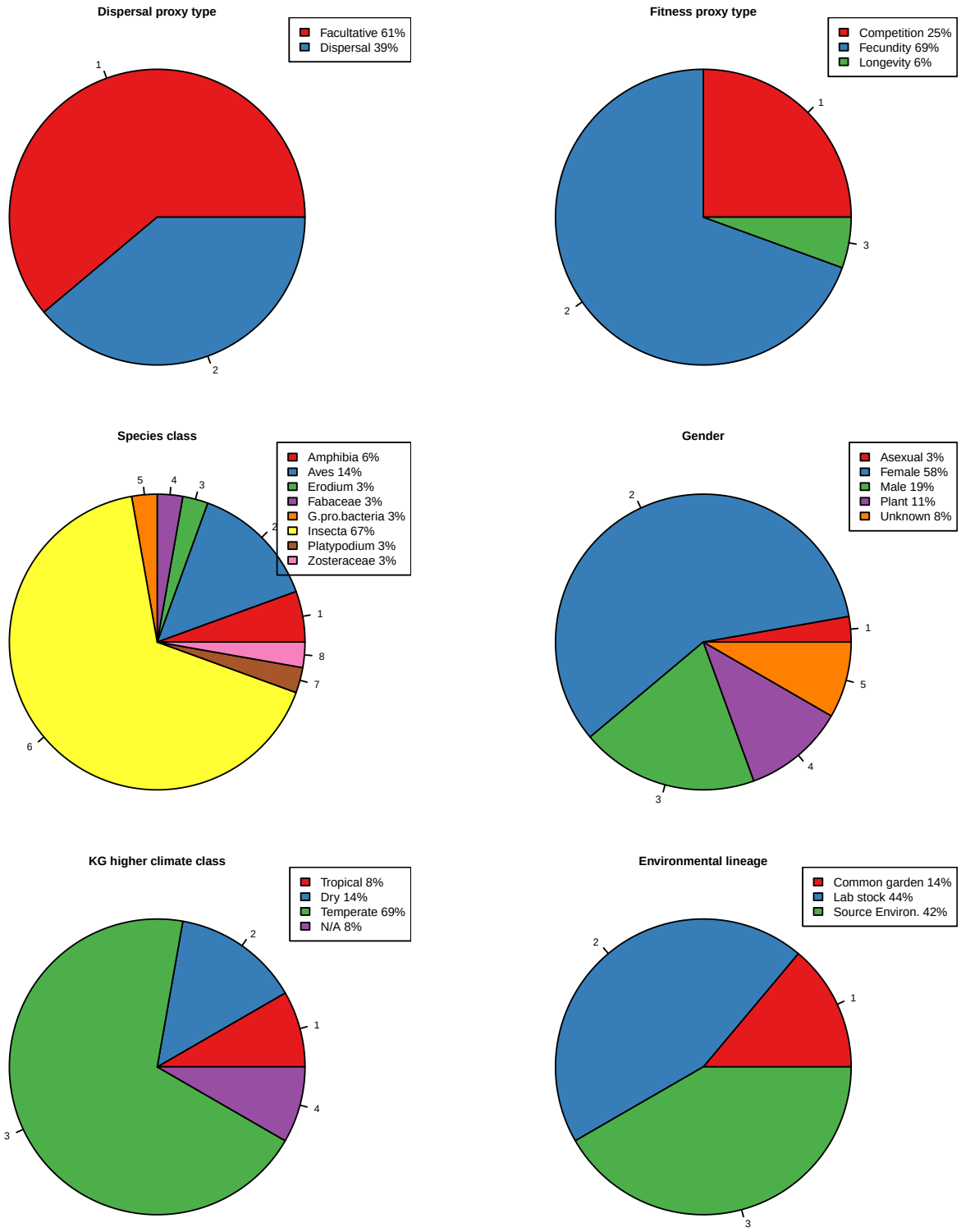

Figure 4.1: Proportions of heterogeneity within classifications of explanatory variables for the continuous data set.

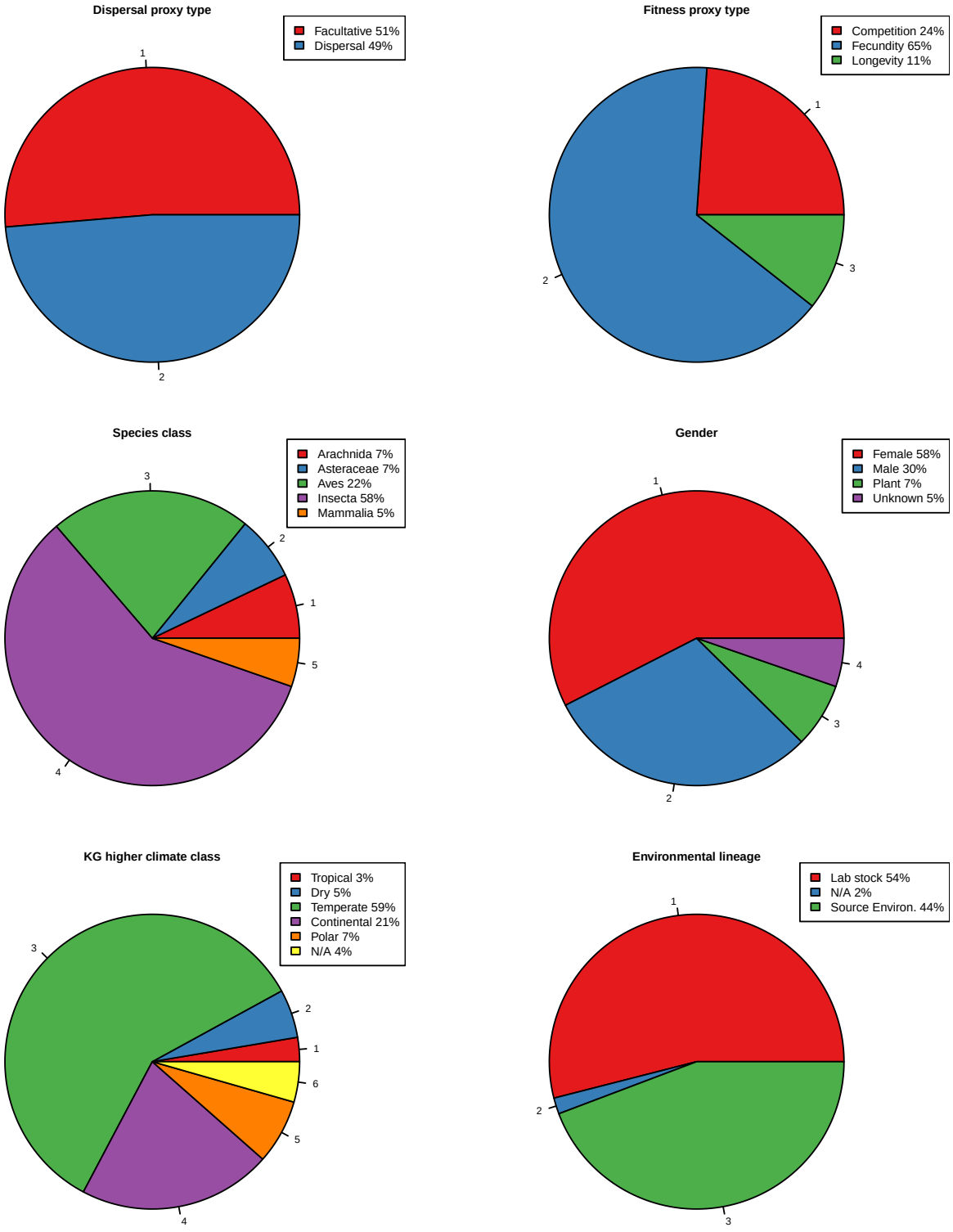

Figure 4.2: Proportions of heterogeneity within classifications of explanatory variables for the binary data set.

4.3 Data synthesis

4.3.1 Effect size calculation

In total we extracted 150 studies from 65 articles for analysis. Effect sizes used in our analysis were calculated in two ways, depending on which of the two data sets they belonged to – Fisher's z for the continuous data, and Hedge's g for the binary data.

Hedge's g is an unbiased standardised mean difference. It is a pooled standardised mean difference with pooled standard deviation between measurements similar to Cohen's d. However, Cohen's d has a bias in the statistic which is resolved by a multiplicative factor – the first product in the relation below – to correct for this which is known as HEdge's g [14]. Hedge's g was calculated from the binary data set by

$$g = \left(1 - \frac{3}{4(n_D + n_{ND}) - 1}\right)\left(\frac{\bar{X}_D - \bar{X}_{ND}}{S_{within}}\right), \tag{4.1}$$

where $\bar{X}_D$ and $\bar{X}_{ND}$ are the sample means of the non-dispersal traits of the dispersers, n_D and n_{ND} are the sample sizes, subscript D, and non-dispersers ND respectively [14]. The within-groups standard deviation pooled across groups S_{within} is given by

$$S_{within} = \sqrt{\frac{(n_D - 1)S_D^2 + (n_{ND} - 1)S_{ND}^2}{(n_D + n_{ND}) - 2}}, \tag{4.2}$$

where S_D and S_{ND} are the standard deviations within the dispersers and non-dispersers respectively. By defining Hedge's g in this way, the convention for a trade-off is that g will be negative. The standard error of g is given by

$$SE_g = \sqrt{\frac{(n_D + n_{ND})}{(n_D n_{ND})} + \left(\left(\frac{\bar{X}_D - \bar{X}_{ND}}{S_{within}}\right)^2 \frac{1}{2(n_D + n_{ND})}\right)} \tag{4.3}$$

The correlation coefficients of the continuous dataset was converted into

Fisher's z statistic for our analysis. The conversion is necessary since the variance for each data point is highly dependent on the correlation coefficient, that is, any correlation coefficient far from 0 with a small sample size will become skewed. This is a problem since within meta-analyses we assume that the sampling distribution is approximately normal. Fisher's z transformation acts as a normalisation step to alleviate this barrier. Once all of the data points are transformed the standard error is given, to a good approximation, by

$$SE_z = \sqrt{\frac{1}{n_z - 3}}, \tag{4.4}$$

where n_z is the sample size of each study [14]. Our original correlations were defined in such a way that a negative correlation indicated a trade-off between dispersal and non-dispersal traits.

Note here that both effect sizes do not scale with the study sample size, and thus can be interpreted as a true biological effect size, rather than a statistical one. This implies any findings regarding the effect sizes have a direct biological interpretation.

Since we found that some articles had multiple studies, we chose to examine the distribution of the calculated effect sizes among articles. The distribution of Fisher's z from the continuous data set can be seen in figure 4.3, and the distribution of Hedge's g from the binary data can be seen in figure 4.4. From figure 4.4 we can see that that within the binary dispersal dataset one article – article 1858 which contained 8 studies, all of which were conducted on the plant species *Asteraceae*, and were the only botanical subjects within the binary data set had a larger absolute effect size relative to the other articles. We believed that this may influence the analysis and not provide an accurate assessment of trade-offs within the dataset. Therefore, we created another dataset without these studies and analysed them separately – one with article 1858 excluded, and the other with 1858 included. We now refer to these data sets as *binary all* when all studies were included, and *binary minus plants* when article 1858 was excluded.

Both analyses were then compared to determine if the exclusion of article 1858 affected the overall trends witnessed when it was included.

Of the binary dispersal data set, 3% of the studies (4 out of 114) were population-level studies and were not separated from those that were at the individual level. Of the continuous dispersal data set, 27% of the studies (10 of the 36) were population level studies and were not separated from those that were at the individual level. This was because the sample sizes within the individual studies that made up the populations were too small to be included as studies. However, those individual studies were able to be aggregated into enough populations to be able to calculate Pearson's correlation coefficient between the populations.

When we conducted analysis regarding both climate and environmental lineage as explanatory variables, we removed any "N/A" values, which accounted for 5 studies within the binary data sets, and 3 studies within the continuous data sets.

4.3.2 Publication bias

To determine if publication bias existed within our dataset we examined enhanced funnel plots for asymmetry – an indication of publication bias – and used Egger's regression test for a quantitative measure of asymmetry within the data. Each study is plotted as a scatter diagram of effect size on the x-axis, and standard error on the y-axis. The funnel shape indicates the quality of the effect size measured for each study with respect to the standard error. Those points that fall within the funnel are more reliable than those outside of it. If there is asymmetry about the central axis – the mean – then publication bias may be present. The mixed-effects model was fitted using the METAFOR package within R (Ver. 3.3.3) [173, 190] and funnel plots were constructed from the linearly predicted models with mixed-random effects. The standard errors used in linear models were calculated using equn. (4.4) for the continuous dispersal data set and equn. (4.3) for the binary dispersal data set.

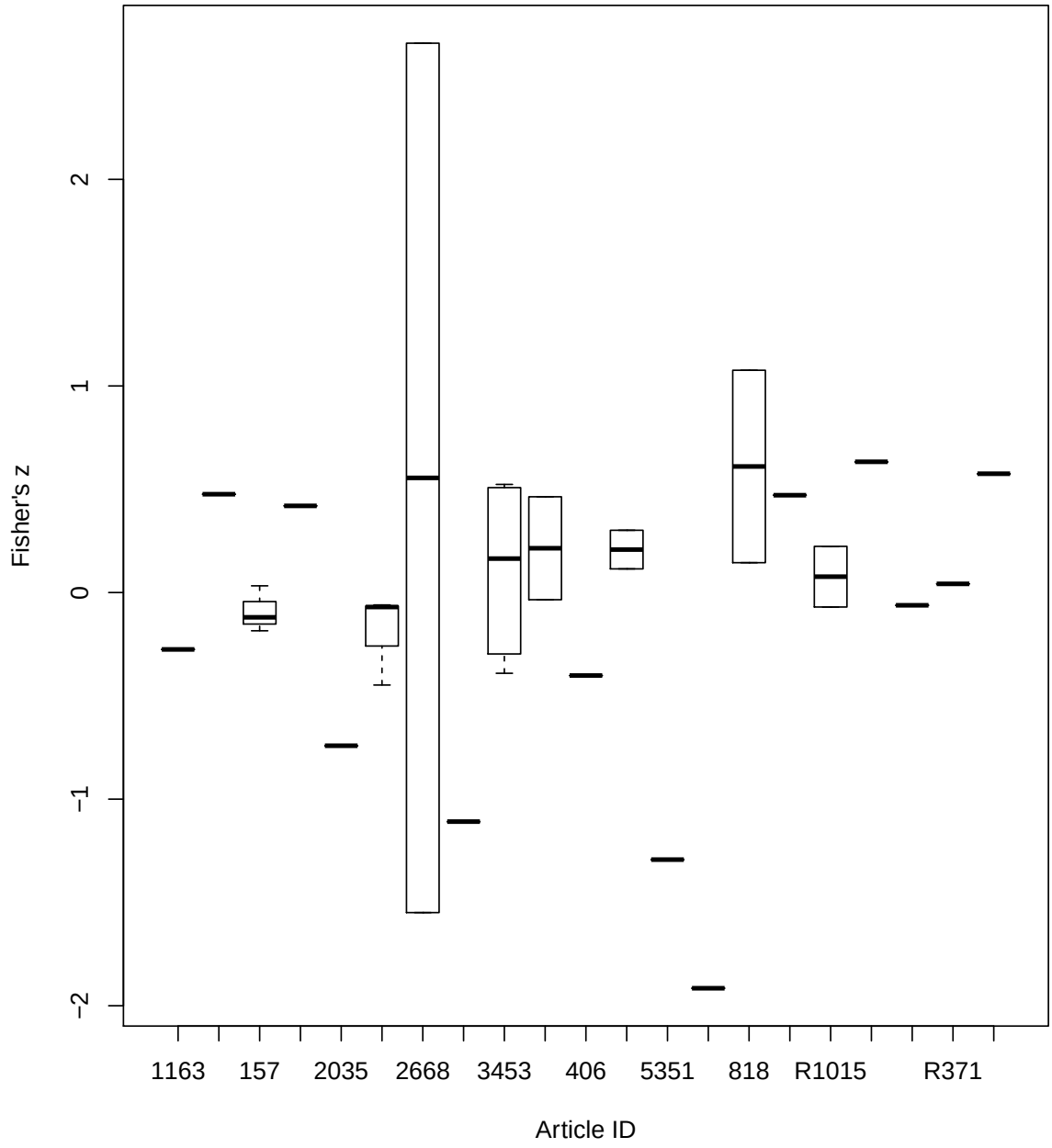

Figure 4.3: Box plots showing the distribution of Fisher's z for studies within articles. Articles that had only one study are shown as means without boxes and whiskers. Biologically this can be interpreted as the range and variation within correlations between dispersal and non-dispersal traits.

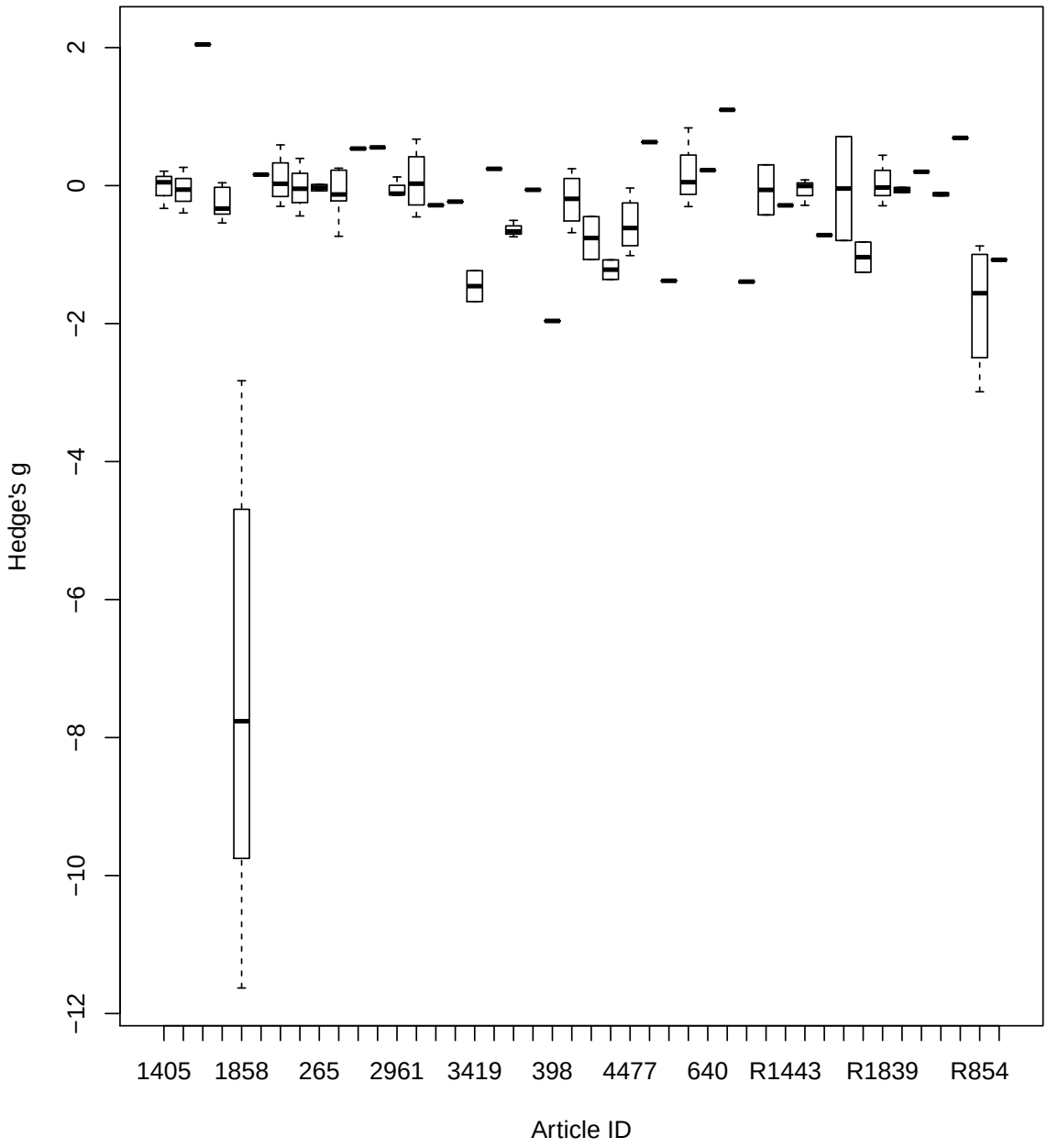

Figure 4.4: Box plots showing the distribution of Hedge's g for studies within articles. Articles that had only one study are shown as means without boxes and whiskers. Article 1858 does not follow the same trend as the remaining articles. Biologically this can be interpreted as the range and variation within the Hedge's g corrected difference between the non-dispersal traits between dispersers and non-dispersers.

Upon inspection of the continuous dispersal data, shown in figure 4.5, there seems to be no asymmetry about the effect size. This was also confirmed by a non-significant z from Egger's regression test (Egger's test: $z = 1.0136$, $p = 0.3108$).

Likewise, upon inspection of the binary dispersal data, shown in figure 4.6, there seems to be no asymmetry about the effect size. We confirmed this in the same way as the correlative data using Egger's regression test for asymmetry (Egger's test: z=-0.6852, p=0.4932)

We should not expect publication bias to be a factor since much of the data collected were from studies where trade-offs were not the sole focus. Therefore, we should expect an unbiased representation of dispersal trade-offs from our samples.

4.4 Statistical analysis

We implemented our meta-analytic approach using the R package MCMCGLMM [70] which fits generalised linear models using a Bayesian framework for estimating effect sizes and variance components [173, 69]. Models within MCMCGLMM accommodate both fixed and random effects meta-analyses while accouting for within-study measurement error variance [69, 198]. We ran each model for $1x10^6$ iterations sampling for the effect sizes and variance every 500^{th} iteration after an initial burn-in of $300,000$. We checked for convergence of MCMC chains by visually inspecting the trace plots, and also ran Gelman-Rubin diagnostics for autocorrelation between samples [60, 197, 172]. Each model was replicated for a total of 3 MCMC chains to ensure that the autocorrelation of each chain was acceptably low to make inferences [60, 197, 172]. An example of diagnostics results are presented within Appendix C.

Effect size precision was accounted for by fitting measurement error variance about the effect size as a variance component, which required the assumption that these were known without error and were fixed [69, 197]. Both the binary all

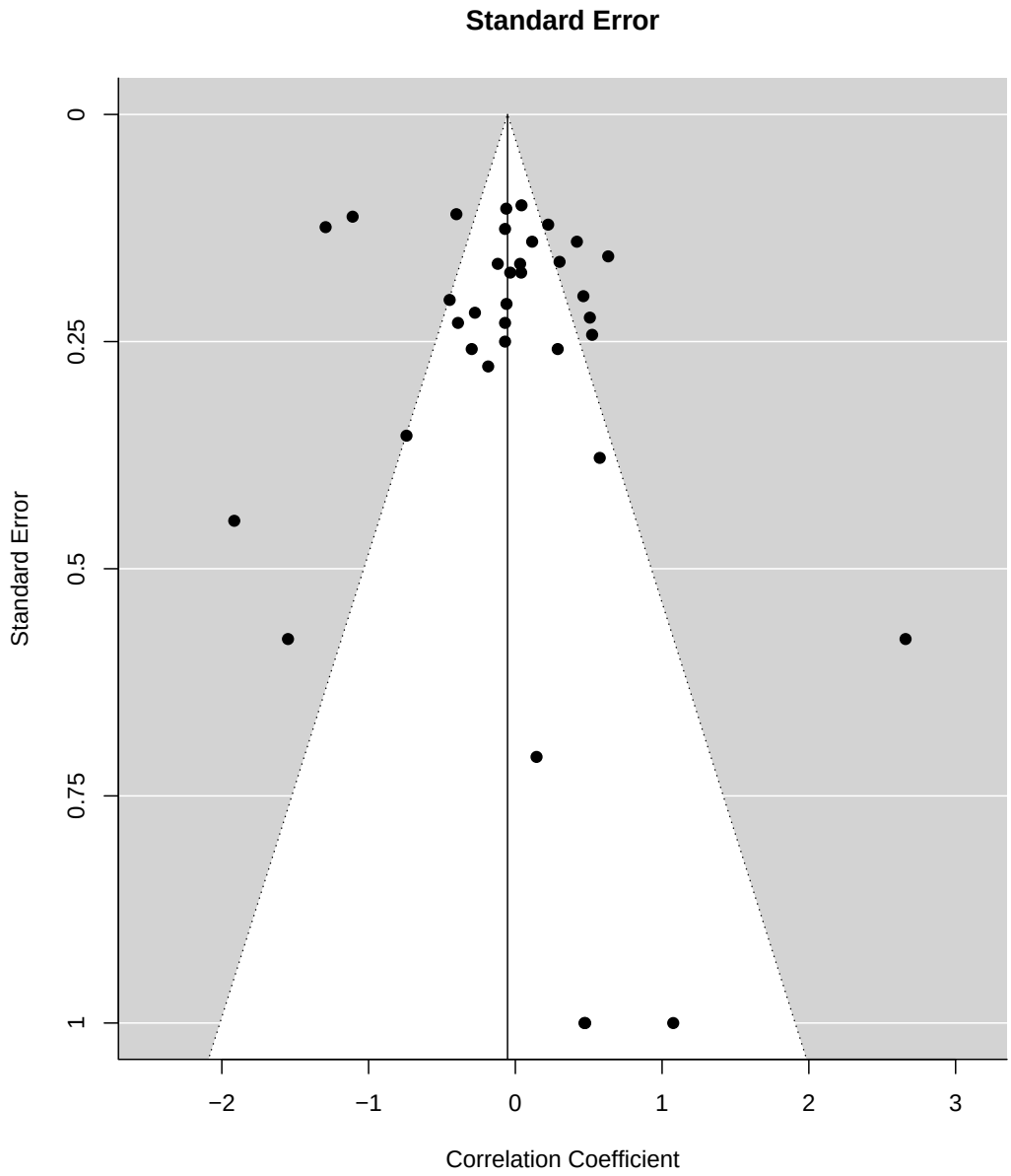

Figure 4.5: Funnel plot of effect size (Fisher's z) vs. standard error of the continuous data set around the effect size. The calcualted overall effect size is shown as the solid vertical line. Data points are approximately symmetrically distributed around the effect size. Studies that are have fewer sample sizes are located at the base of the funnel.

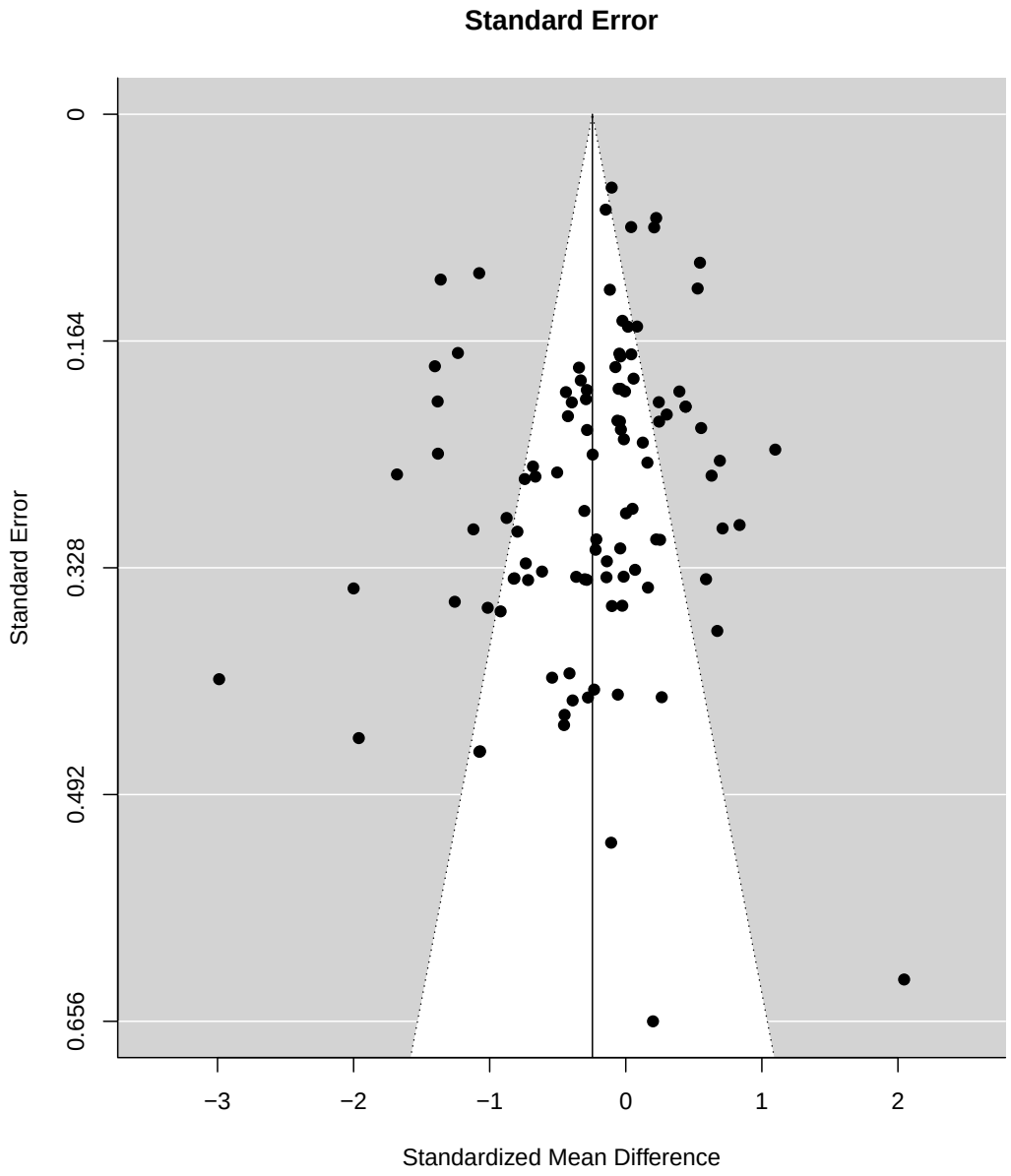

Figure 4.6: Funnel plot of effect size (standardised mean difference) vs. standard error of the binary all data set around the effect size. The calculated overall effect size is shown as the solid vertical line. Data points are approximately symmetrically distributed around the effect size. Studies that are have fewer sample sizes are located at the base of the funnel.

and binary minus plants data sets had the article from which they were taken fit as random effects, since it is expected that multiple studies from the same article will have similar effect sizes. Random effects were not fitted for the continuous dispersal data set because the model could not distinguish between residual and a random effect for the articles.

Our explanatory variables identified in section 4.2.4 were fitted as fixed effects. We developed hypotheses to test whether dispersal trade-offs were present from the literature collected, and how strong any trade-off is. Furthermore, we also wanted to determine how the explanatory variables influenced the expression of a dispersal trade-off. We did not want to make any assumptions about our data before our analysis and thus selected to use non-informative improper priors for both fixed and random effects (corresponding to $V = 1\text{x}10^{-16}, \nu = -1$ within MCMCGLMM) [59].

The models will determine the estimate of the pooled effect size and whether this is different from zero. If the calculated 95% - credible intervals (CIs) do not significantly span 0, we conclude that the effect size is different from zero. Note that this is how we determined if the effect size was significant or not when referring to *significance* by this definition. If the pooled effect size is less than zero and CIs as previously described, then we can conclude that a trade-off exists.

When reporting results, we specify that a result with a negative effect size is significant when the CIs do not span 0, and insignificant if they do. The package MCMCGLMM reports a statistic called *pMCMC* with indicators familiar in p-value testing, however it is not a true p-value. The pMCMC value is an indicator of significance of whether CIs span 0, it is calculated as two times the smaller of two quantities: the probability that the effect size is less than zero, and the probability that it is greater than 0.

Firstly, we fitted a model to all of the data sets which accounts for the average pooled effect of all studies, that is, with only the intercept fitted. This

was repeated on the binary data sets with random effects for a total of 5 models. We then followed the previous format for each of our explanatory variables and removed the intercept to determine which of the parameters best explained any effect the first models produced.

We then compared the models with and without random effects fitted to both binary data sets using the reported *deviance information criterion* (DIC) which is a hierarchical modelling method for determining which model is the best fit to the data [169]. The DIC value is measure of how well the models fitted the data whereby the lowest value is considered the best fit.

We have previously mentioned that random effects were fitted to account for the heterogeneity between articles. Variance within effect sizes that are attributed to hierarchical factors inherent within meta-analyses can be accounted for proportionally to provide deeper understanding of any results [172]. The total variance for our analysis can be defined by

$$\sigma_T^2 = \sigma_{ArtID}^2 + \sigma_E^2 + \sigma_M^2, \tag{4.5}$$

where σ_T^2 is total variance, σ_{ArtID}^2 is the variance component attributed to article ID, σ_E^2 is the component attributed to error variance (residuals), and σ_T is the 'typical' measurement error variance which is given by

$$\sigma_M^2 = \frac{\sum_{i=1}^{k} w_i(k-1)}{\left(\sum_{i=1}^{k} w_i\right)^2 - \sum_{i=1}^{k} w_i^2}, \tag{4.6}$$

where w_i is the inverse of the i^{th} measurement error variance associated with study i [172, 80].

4.5 Meta-analytic Results

We report the results of each model fitted to the three datasets and whether a dispersal trade-off was found. A visual representation of our results are depicted

within figures and the full results of the models described within tables. The figures are presented as the estimated mean effect size with CIs for that estimate. The entries of the table present the model and data sets which they were fitted to, the DIC value for that model, the parameter of interest, the estimated mean effect size and CIs, the proportion of variance attributed to random effects, the effective sample sizes from which the mean was calculated from, and the pMCMC value.

The results from the continuous data set were all non-significant with the exception of the competitive parameter, which will be discussed within that section. Since the results were non-significant, we leave any interpretation until the discussion. Thus, when referring to models and data sets within this section (section 4.5), *"all"* refers exclusively to the binary data sets only.

We found that models which included random effects for article ID that were fitted to the binary datasets had a consistently lower DIC than those without. We can then conclude that accounting for random effects between articles are better models.

4.5.1 Overall effect

All models fitted to all data sets were estimated to have negative posterior mean coefficients with CIs that did not significantly span 0 (fig. 4.7, table 4.3). Including random effects weakened the effect size and narrowed the CIs. Therefore, evidence of an overall dispersal trade-off was present.

Table 4.3: Mixed effect models for each dataset for an overall effect of a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.7. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var.** σ^2_{ArtID}	**eff. samp.**	**pMCMC**
● cont.	z ~1	intercept	69.157	-0.052	(-0.252, 0.170)	0	1400	0.629
● bin. all	g ~1	intercept	453.168	-0.676	(-0.993, -0.347)	0	1683	$< 6\text{x}10^{-4}$ ***
● bin. all	g ~1 + random(Article ID)	intercept	218.763	-0.377	(-0.731, -0.001)	80.92	1800	0.041 *
● bin. - plants	g ~1	intercept	205.592	-0.244	(-0.363, -0.116)	0	1835	0.001 **
● bin. - plants	g ~1 + random(Article ID)	intercept	52.915	-0.221	(-0.419, -0.001)	79.69	1800	0.044 *

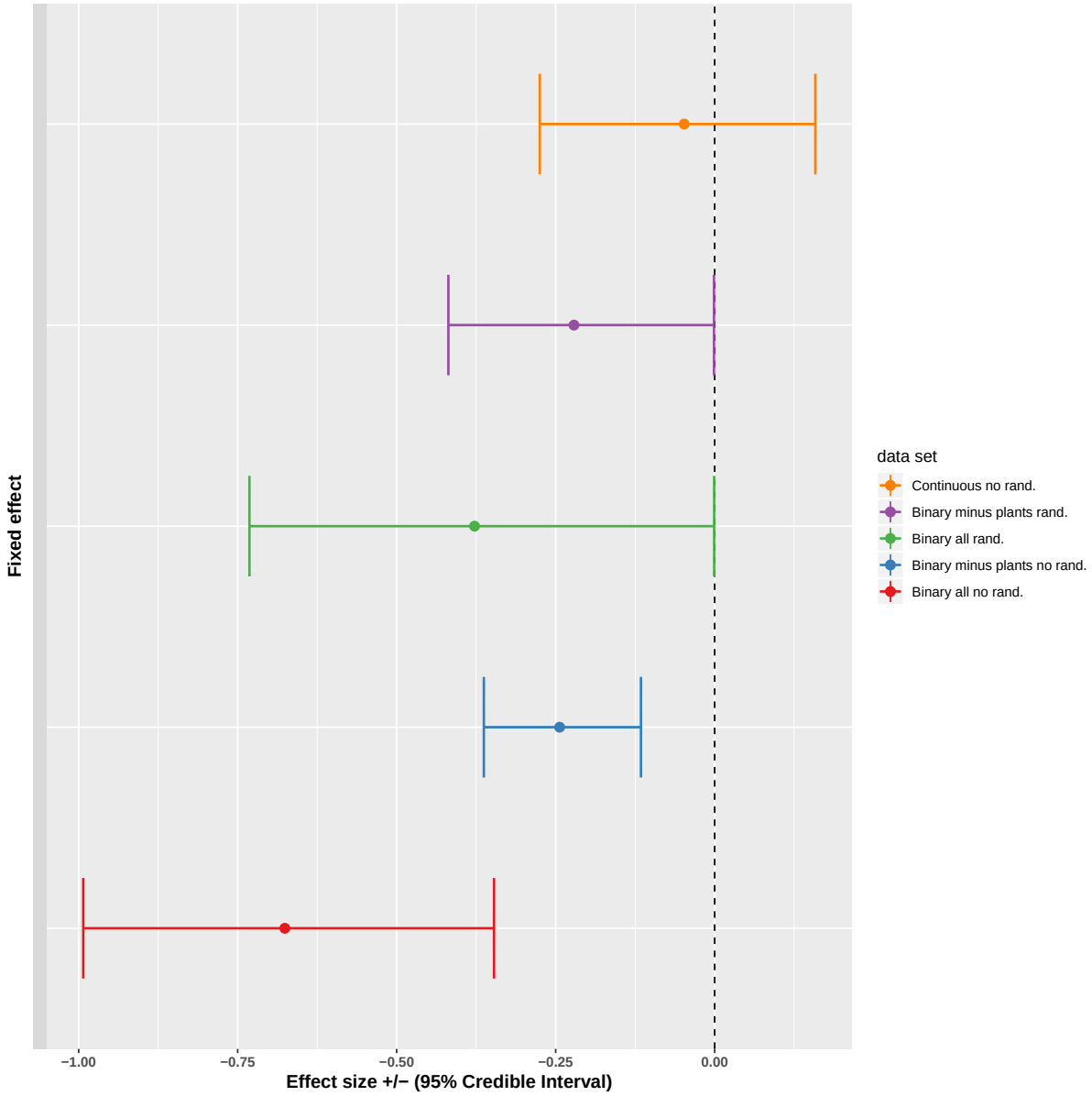

Figure 4.7: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.3. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off.

4.5.2 Effect of dispersal type

Facultative investment in dispersal

Models fitted to the binary minus plants data set were estimated to have positive posterior mean coefficients, while estimate from the binary full data set were negative (fig. 4.8, table 4.4). All models estimated CIs that significantly spanned 0, indicating that there is no evidence for a trade-off between facultative dispersal and non-dispersal traits.

Physical investment in dispersal

All models fitted to all data sets were estimated to have negative posterior mean coefficients with CIs that did not significantly span 0 (fig. 4.8, table 4.4). Including random effects weakened the effect and narrowed the CIs. Therefore, there is evidence of a dispersal trade-off between physical investment in dispersal and non-dispersal traits.

Table 4.4: Mixed effect models for each dataset for an effect of dispersal type within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.8. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var.** σ^2_{ArtID}	**eff. samp.**	**pMCMC**
● cont.	z ~disp. type - 1	Facultative	69.548	0.004	(-0.248, 0.243)	0	1400	0.954
		Phys. Invest.		-0.178	(-0.547, 0.228)		1400	0.356
● bin. all	g ~disp. type - 1	Facultative	437.420	-0.053	(-0.446, 0.397)	0	1800	0.822
		Phys. Invest.		-1.337	(-1.767, -0.874)		1800	$< 6\text{x}10^{-4} * **$
● bin. all	g ~disp. type + random(Article ID) - 1	Facultative	213.958	0.079	(-0.446, 0.492)	78.86	1900	0.737
		Phys. Invest.		-0.854	(-1.350, -0.386)		1800	0.001**
● bin. - plants	g ~disp. type - 1	Facultative	194.534	-0.056	(-0.220, -0.094)	0	1981	0.489
		Phys. Invest.		-0.483	(-0.665, -0.318)		2085	$< 6\text{x}10^{-4} * **$
● bin. - plants	g ~disp. type + random(Article ID) - 1	Facultative	47.695	0.065	(-0.185, 0.320)	76.51	1800	0.616
		Phys. Invest.		-0.543	(-0.817, -0.267)		1800	$< 6\text{x}10^{-4} * **$

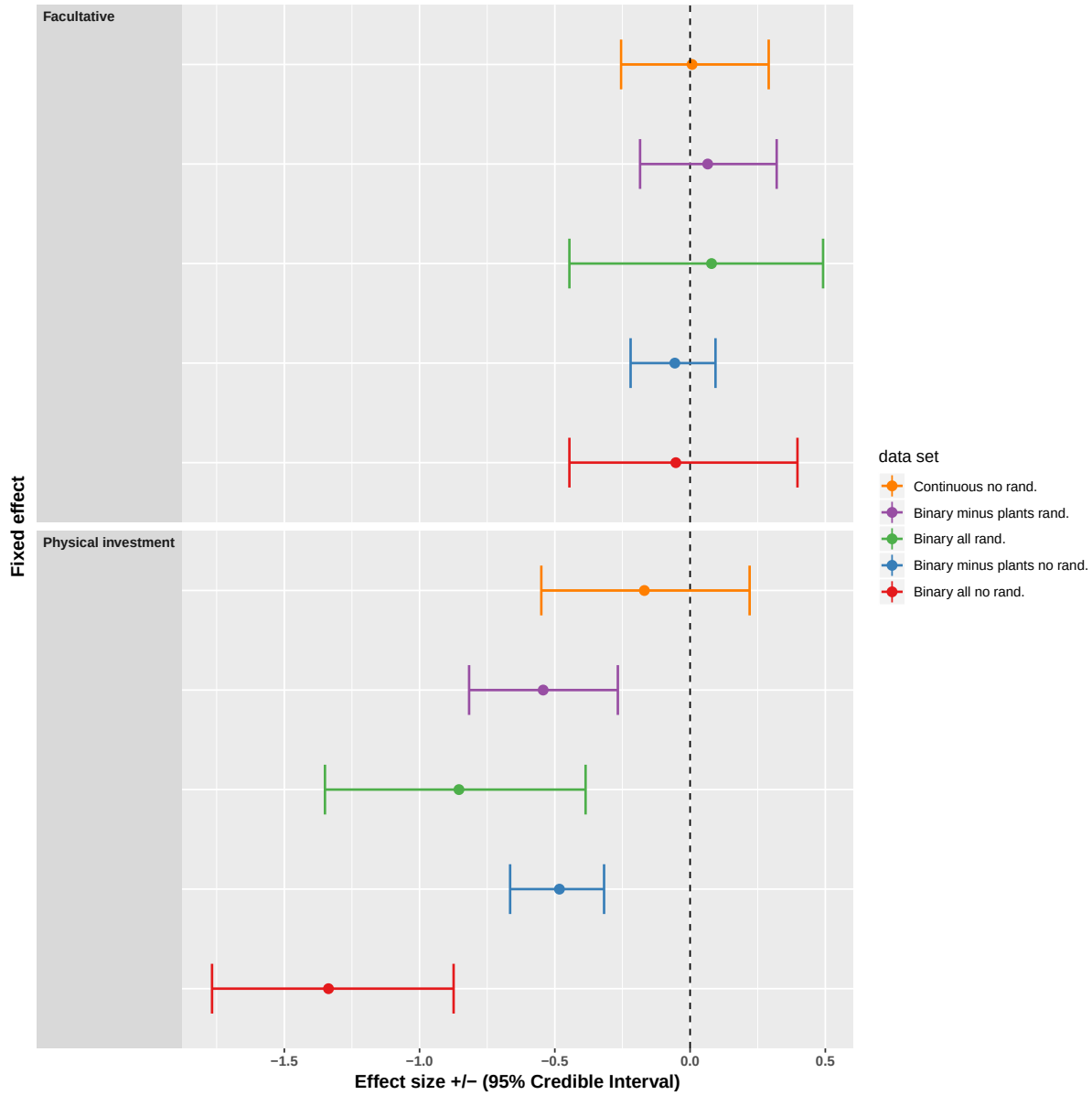

Figure 4.8: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.4. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.5.3 Effect of fitness type

Results from the models are summarised in table 4.5, and figure 4.9.

Competitive

All models fitted to all data sets were estimated to have negative posterior mean coefficients (fig.4.9, and table 4.5). Both models that were fitted to the binary minus plants data set had non-significant CIs. The model fitted to the binary full data set with no random effects had significant CIs, but became non-significant when random effects were included within the model. The continuous data set had significant CIs. The inclusion of random effects weakened the effects and narrowed the CIs. There is then little evidence for a dispersal trade-off with competitive fitness ability.

Fecundity

All of the models estimated negative posterior mean coefficients (fig. 4.9 and table 4.5). CIs from the models fitted to both binary data sets were non-significant when random effects were excluded, but became significant when random effects were included.

Longevity

All of the models estimated negative posterior mean coefficients with non-significant CIs (fig. 4.9 and table 4.5). Inclusion of random effects weakened the effect within the binary minus plants, and had no effect on the estimate from the binary all data. The CIs were also narrower when random effects were included. There is then no evidence for a dispersal trade-off with longevity.

Table 4.5: Mixed effect models for each dataset for an effect of fitness type within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.9. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var.** σ^2_{ArtID}	**eff. samp.**	**pMCMC**
● cont.	z ~fitness - 1	Competitive	66.050	-0.448	(-0.861, -0.039)	0	1400	0.034*
		Fecundity		0.101	(-0.142, 0.347)		1400	0.413
		Longevity		-0.046	(-0.805, 0.784)		1132	0.883
● bin. all	g ~fitness - 1	Competitive	439.894	-1.909	(-2.543, -1.260)	0	1972	$< 6\text{x}10^{-4}$ ***
		Fecundity		-0.328	(-0.703, 0.069)		1800	0.097 .
		Longevity		-0.081	(-1.091, 0.804)		1543	0.828
● bin. all	g ~fitness + random(Article ID) - 1	Competitive	227.656	-0.477	(-1.057, 0.070)	79.69	1800	0.092 .
		Fecundity		-0.352	(-0.745, 0.024)		1800	0.074 .
		Longevity		-0.238	(-0.771, 0.330)		1800	0.411
● bin. - plants	g ~fitness - 1	Competitive	205.444	-0.064	(-0.331 ,0.238)	0	1800	0.642
		Fecundity		-0.321	(-0.464, -0.178)		1800	$< 6\text{x}10^{-4}$ ***
		Longevity		-0.070	(-0.438, 0.291)		1752	0.734
● bin. - plants	g ~fitness + random(Article ID) - 1	Competitive	56.055	-0.120	(-0.467, 0.172)	79.72	1800	0.450
		Fecundity		-0.265	(-0.470, -0.034)		1611	0.021*
		Longevity		-0.095	(-0.384, 0.243)		1800	0.574

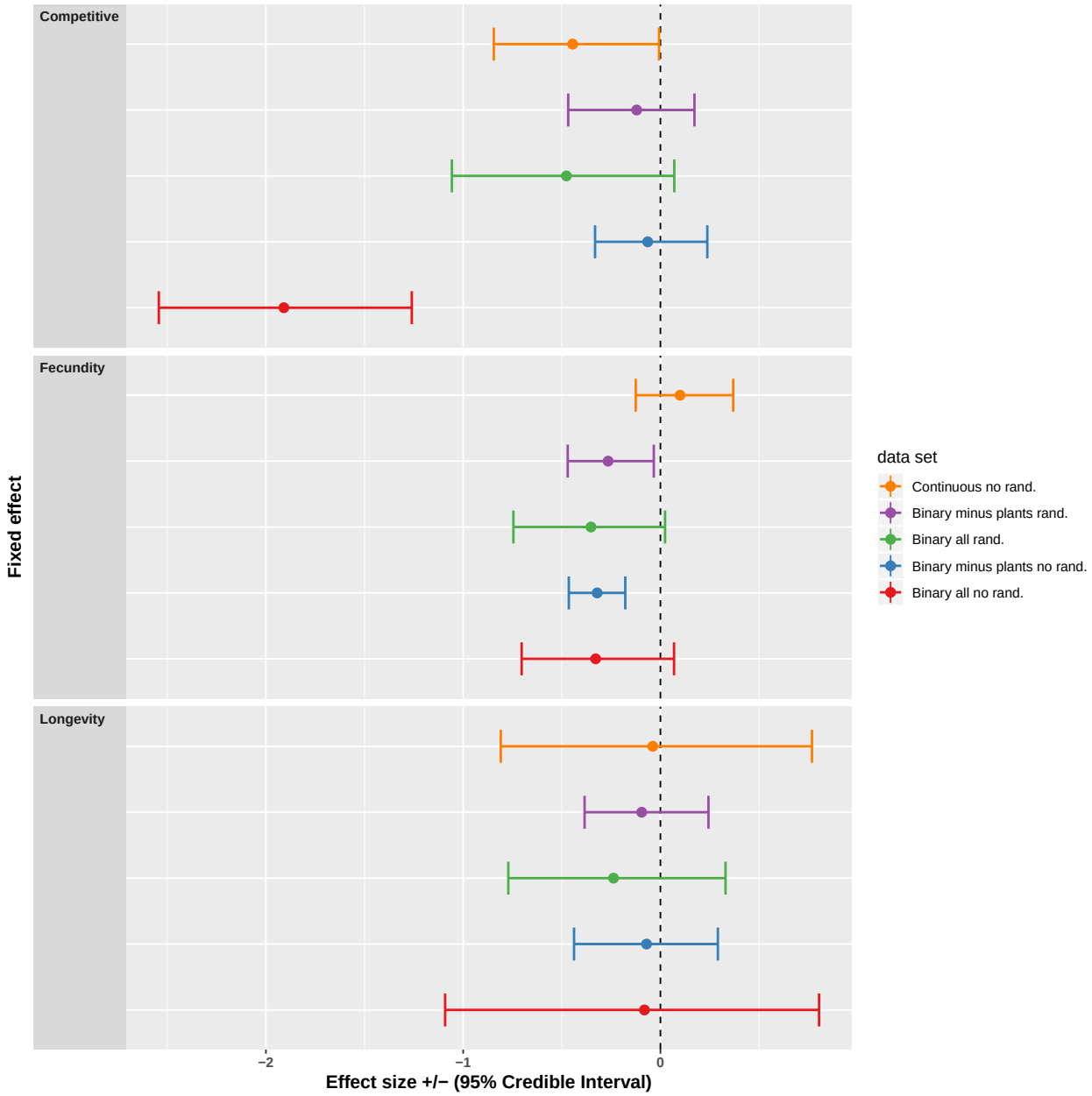

Figure 4.9: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.5. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.5.4 Effect of sex

Female

All of the models estimated negative posterior mean coefficients all of which had significant CIs (fig. 4.10 and table 4.6). The CIs of the model with random effects fitted to the binary minus plants data had marginally significant CIs. The inclusion of random effects had no strong noticeable effect on either the estimated posterior mean, nor the CIs. There is then evidence of a dispersal trade-off among female subjects.

Male

All of the models estimated negative posterior mean coefficients, all of which had non-significant CIs (fig. 4.10 and table 4.6). The inclusion of random effects had no strong noticeable effect on either the estimated posterior mean, nor the CIs. There is then no evidence of a dispersal trade-off among male subjects.

Plant

All of the models estimated negative posterior mean coefficients all of which had significant CIs (fig. 4.10 and table 4.6). Models were fitted to the binary all and continuous data sets. The effect was stronger in the binary all data set than the continuous data set. The inclusion of random effects widened the CIs. Therefore, there is evidence of a dispersal trade-off within plants.

Unknown

All of the models estimated negative posterior mean coefficients all of which had non-significant CIs (fig. 4.10 and table 4.6). When random effects were fitted the effect size was stronger with narrower CIs within the binary minus plants data set. There was no noticeable change within effect size of the binary full data set, but the CIs were narrower. There is no evidence of a dispersal trade-off when sex was unknown.

Table 4.6: Mixed effect models for each dataset for an effect of sex within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.10. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var.** σ^2_{ArtID}	**eff. samp.**	**pMCMC**
● cont.	z ~Sex - 1	Asexual	70.357	-0.090	(-1.303, 0.955)	0	1285	0.867
		Female		0.079	(-0.188, 0.371)		1400	0.567
		Male		0.015	(-0.516, 0.459)		1295	0.931
		Plant		-0.613	(-1.209 -0.060)		1400	0.026 *
		Unknown		-0.160	(-0.839 0.638)		1400	0.650
● bin. all	g ~Sex - 1	Female	278.286	-0.313	(-0.518, -0.117)	0	1800	0.006**
		Male		-0.1494	(-0.421, 0.118)		2028	0.274
		Plant		-6.502	(-7.2285, -5.7601)		1800	$< 6\text{x}10^{-4}$ ***
		Unknown		-0.117	(-0.728, 0.560)		1800	0.718
● bin. all	g ~Sex + random(Article ID) - 1	Female	252.890	-0.296	(-0.546, -0.078)	25.95	1800	0.016*
		Male		-0.162	(-0.460, 0.164)		1438	0.300
		Plant		-6.361	(-7.418, -5.243)		1800	$< 6\text{x}10^{-4}$ ***
		Unknown		-0.213	(-0.945, 0.582)		1610	0.586
● bin. - plants	g ~Sex - 1	Female	207.053	-0.311	(-0.492, -0.155)	0	1800	$< 6\text{x}10^{-4}$ ***
		Male		-0.144	(-0.372, 0.064)		1799	0.220
		Unknown		-0.126	(-0.610, 0.376)		1800	0.577
● bin. - plants	g ~Sex + random(Article ID) - 1	Female	58.487	-0.213	(-0.452, 0.009)	79.88	1800	0.063 .
		Male		-0.198	(-0.452, 0.069)		2028	0.148
		Unknown		-0.411	(-1.070, 0.277)		1800	0.218

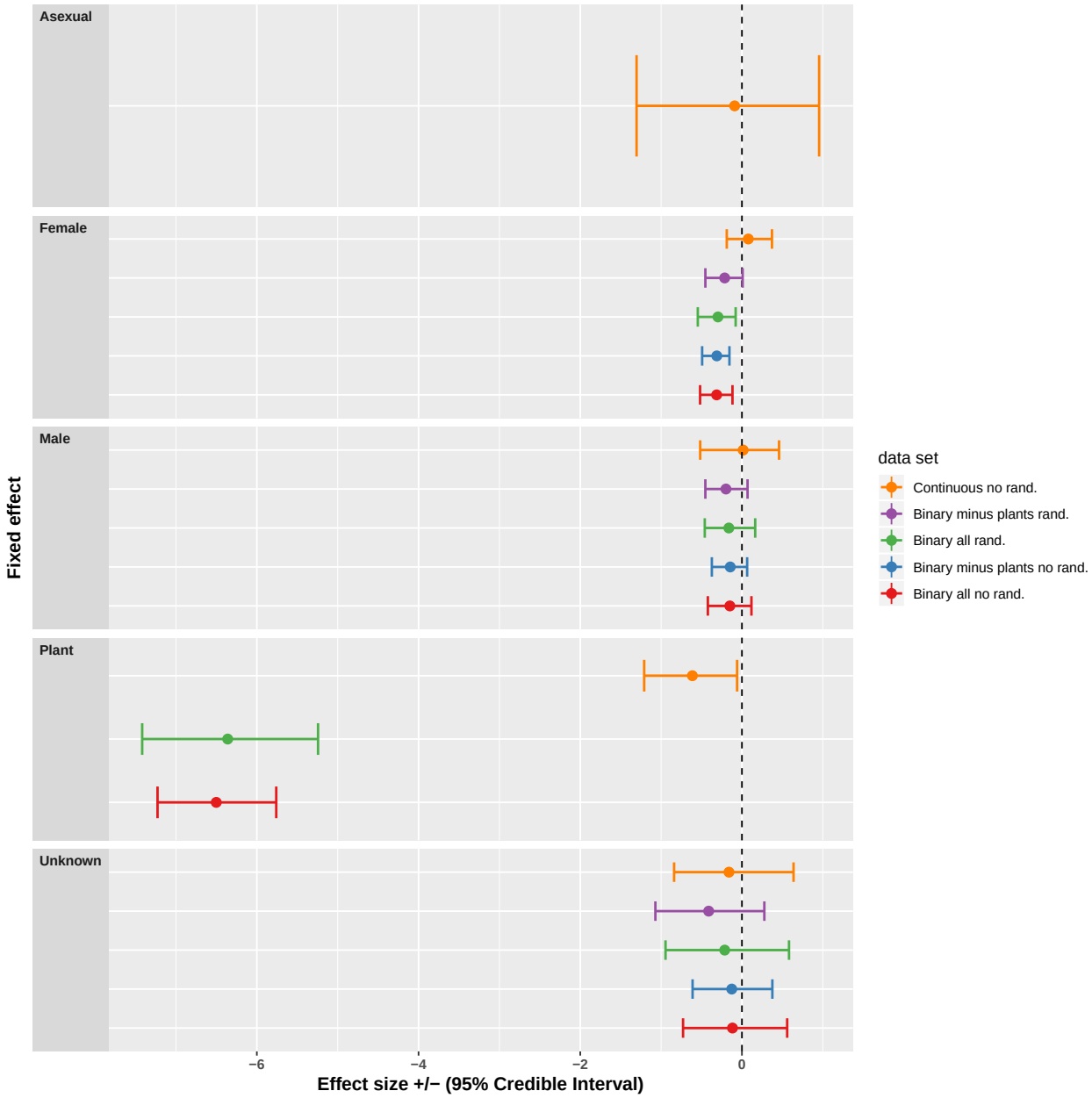

Figure 4.10: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.6. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.5.5 Effect of species class

Arachnida

All of the models estimated negative posterior mean coefficients all of which had non-significant CIs, with one exception of of the binary minus data with no random effects which were narrowly significant (fig. 4.11 and table 4.7). The CIs of the binary all data with random effect fitted were narrower than without, whereas the remainder had no strong noticeable effect on either the estimated posterior mean, nor CIs. There is then no strong evidence of a dispersal trade-off within Arachnida.

Aves

All of the models estimated negative posterior mean coefficients all of which had non-significant CIs (fig. 4.11 and table 4.7). The CIs of the binary all data with random effects fitted were narrower than without, whereas the CIs of the binary minus plants were wider when random effects were included. There is then no strong evidence of a dispersal trade-off within Aves.

Insecta

All of the models estimated negative posterior mean coefficients all of which had significant CIs (fig. 4.11 and table 4.7). The inclusion of random effects had no strong noticeable effect on either the estimated posterior mean, nor the CIs. There is then evidence of a dispersal trade-off within Insecta.

Mammalia

All of the models estimated positive posterior mean coefficients, all of which had non-significant CIs (fig. 4.11 and table 4.7). The inclusion of random effects had no strong noticeable effect on either the estimated posterior mean, nor the CIs. There is then no evidence of a dispersal trade-off within Mammalia.

Asteraceae

Both estimates of the posterior mean coefficients were strongly negative with highly significant CIs (fig. 4.11 and table 4.7). The inclusion of random effects widened the CIs of the estimate. There is then evidence for a dispersal trade-off within Asteraceae.

Table 4.7: Mixed effect models for each dataset for an effect of species class within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.11. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are Signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var. σ^2_{ArtID}**	**eff. samp.**	**pMCMC**
● cont.	z ~Spec. class - 1	Amphibia	53.205	-0.475	(-1.238, 0.227)	0	1400	0.186
		Aves		0.196	(-0.203, 0.654)		1652	0.363
		Erodium		0.422	(-0.511, 1.398)		1400	0.357
		Fabaceae		-1.294	(-2.231, -0.365)		1400	0.010 **
		G.pro.bacteria		-0.057	(-0.925, 0.854)		1400	0.879
		Insecta		-0.036	(-0.180, 0.250)		1400	0.727
		Platypodium		-0.378	(-1.441, 0.471)		1296	0.393
		Zosteraceae		-1.101	(-2.043, -0.124)		1400	0.031
● bin. disp. all	g ~Spec. class - 1	Arachnida	269.344	-0.435	(-0.940, 0.113)	0	1800	0.127
		Asteraceae		-6.452	(-7.166, -5.755)		1800	$< 6x10^{-4}$ ***
		Aves		-0.016	(-0.316, 0.286)		1800	0.926
		Insecta		-0.380	(-0.565, -0.181)		1800	$< 6x10^{-4}$ ***
		Mammalia		0.508	(-0.097, 1.200)		1800	0.119
● bin. disp. all	g ~Spec. class + random(Article ID) - 1	Arachnida	51.862	-0.166	(-1.0951, 0.655)	77.4	1800	0.690
		Asteraceae		-6.288	(-7.224, -5.134)		1800	$< 6x10^{-4}$ ***
		Aves		-0.026	(-0.448, 0.389)		2096	0.928
		Insecta		-0.395	(-0.613, -0.172)		1800	$< 6x10^{-4}$ ***
		Mammalia		0.521	(-0.113, 1.134)		2042	0.111
● bin. disp. -1858	g ~Spec. class - 1	Arachnida	196.695	-0.424	(-0.896, -0.0003)	0	1800	0.0567 .
		Aves		-0.020	(-0.262, 0.213)		1800	0.853
		Insecta		-0.378	(-0.533, -0.232)		1800	$< 6x10^{-4}$ ***
		Mammalia		0.496	(-0.049, 1.007)		1800	0.070 .
● bin. - plants	g ~Spec. class + random(Article ID) - 1	Arachnida	51.862	-0.353	(-1.162, 0.396)	77.97	1961	0.357
		Aves		-0.009	(-0.502, 0.493)		1800	0.974
		Insecta		-0.403	(-0.622, -0.154)		1800	0.002 **
		Mammalia		0.498	(-0.056, 1.073)		1800	0.086 .

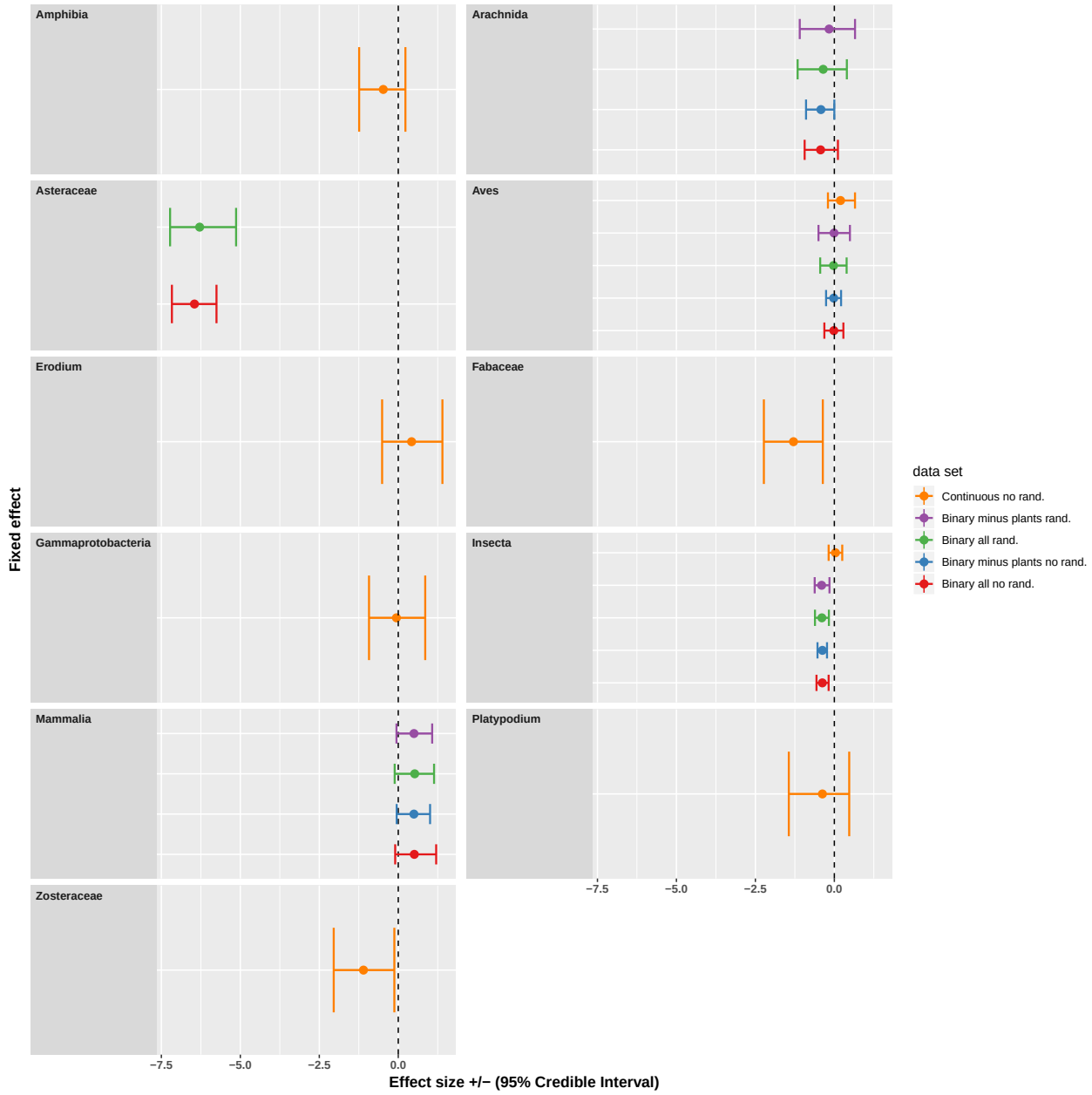

Figure 4.11: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.7. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.5.6 Effect of Köppen-Geiger climate class

Tropical

All of the models estimated positive posterior mean coefficients all of which had non-significant CIs, with one exception (fig. 4.12 and table 4.8). The CIs of the model with random effects fitted to the binary minus plants data had marginally significant CIs. The inclusion of random effects weakened the effect and narrowed the CIs. There is then little evidence of a dispersal trade-off among subjects from tropical climates.

Dry

All of the models estimated positive posterior mean coefficients all of which had non-significant CIs (fig. 4.12 and table 4.8). The inclusion of random effects weakened the effect and narrowed the CIs. There is then no evidence of a dispersal trade-off among subjects from dry climates.

Temperate

All of the models estimated negative posterior mean coefficients all of which had significant CIs (fig. 4.12 and table 4.8). The inclusion of random effects weakened the effect and narrowed the CIs. There is then evidence of a dispersal trade-off among subjects from temperate climates.

Continental

All of the models estimated positive posterior mean coefficients all of which had non-significant CIs (fig. 4.12 and table 4.8). The inclusion of random effects weakened the effect and narrowed the CIs. There is then no evidence of a dispersal trade-off among subjects from continental climates.

Polar

All of the models estimated positive posterior mean coefficients all of which had non-significant CIs (fig. 4.12 and table 4.8). The inclusion of random effects had no effect on the estimate of the posterior mean, but narrowed the CIs. There is

then no evidence of a dispersal trade-off among subjects from polar climates.

Table 4.8: Mixed effect models for each dataset for an effect of climate within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.12. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	Model	Parameter	DIC	Estimated mean	95% CI intervals	% of Var. σ^2_{ArtID}	eff. samp.	pMCMC
● cont.	z ~climate - 1	Tropical	55.497	-0.443	(-1.079, 0.1505	0	1400	0.166
		Dry		-0.418	(-0.876, 0.073)		1400	0.090
		Temperate		0.113	(-0.108, 0.348)		1400	0.320
● bin. all	g ~climate - 1	Tropical	438.494	0.986	(-1.252, 2.942)	0	1800	0.357
		Dry		0.269	(-1.219, 1.633)		1800	0.724
		Temperate		-1.043	(-1.483, -0.613)		1800	$< 6x10^{-4}$ ***
		Continental		-0.347	(-1.034, 0.372)		1800	0.340
		Polar		-0.285	(-1.666, 0.889)		1800	0.642
● bin. all	g ~climate + random(Article ID) - 1	Tropical	221.145	0.934	(-0.617, 2.377)	79.62	1800	0.227
		Dry		0.498	(-1.086, 1.850)		1768	0.494
		Temperate		-0.580	(-1.053, -0.0120)		1800	0.012*
		Continental		-0.524	(-1.347, 0.169)		1800	0.161
		Polar		-0.298	(-1.663, 1.161)		1800	0.650
● bin. - plants	g ~climate - 1	Tropical	177.457	0.740	(-0.054, 1.367)	0	1800	0.0367 *
		Dry		0.238	(-0.188, 0.717)		1800	0.299
		Temperate		-0.408	(-0.561, -0.257)		1673	$< 6x10^{-4}$ ***
		Continental		-0.051	(-0.273, 0.202)		1800	0.680
		Polar		-0.247	(-0.626, 0.120)		1800	0.189
● bin. - plants	g ~climate + random(Article ID) - 1	Tropical	65.245	0.781	(-0.003, 1.606)	72.52	1800	0.058.
		Dry		0.4667	(-0.242, 1.133)		1800	0.172
		Temperate		-0.406	(-0.652, -0.179)		1800	0.001**
		Continental		-0.151	(-0.569, 0.2456)		1800	0.494
		Polar		-0.297	(-0.943, 0.387)		1800	0.368

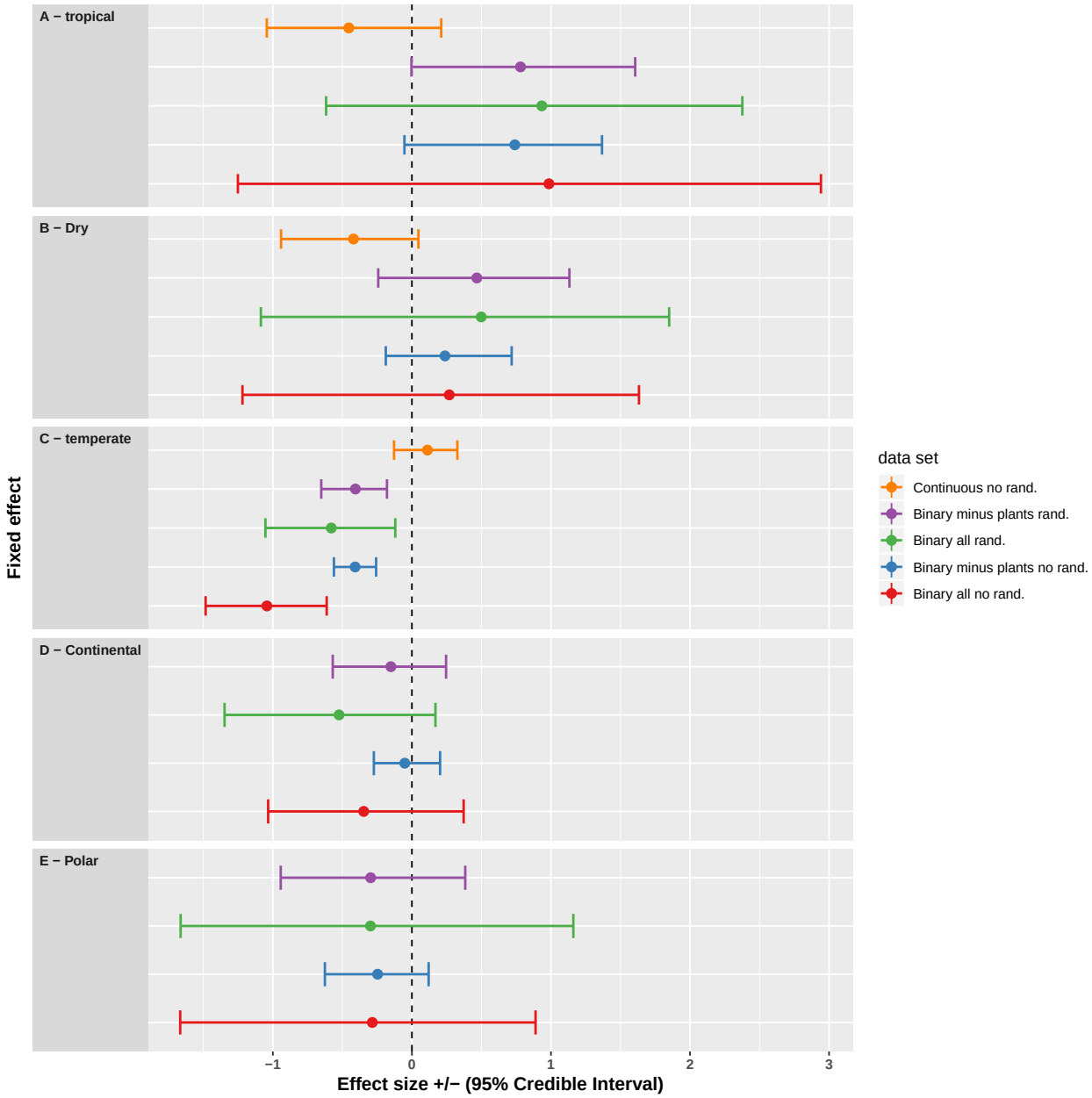

Figure 4.12: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.8. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.5.7 Effect of environmental lineage

Laboratory stock

All of the models estimated negative posterior mean coefficients all of which had non-significant CIs, with one exception (fig. 4.13 and table 4.9). The binary all data set had significant CIs when random effects were included, but this effect disappeared within the binary minus plants data set. When random effects were fitted the estimated posterior means were similar, however the CIs were narrower. There is then no strong evidence of a dispersal trade-off among subjects reared within laboratory stocks.

Sourced from natural environment

All of the models estimated negative posterior mean coefficients (fig. 4.13 and table 4.9). The binary all data set had significant CIs, however this effect disappeared within the binary minus plants data set. The inclusion of random effects weakened the effect and narrowed the CIs. There is therefore no strong evidence for a dispersal trade-off within organisms from natural environments.

We should also note that within the binary minus plants data set the models fitted that included random effects show that there is no noticeable difference between the estimates of the posterior means and CIs of "Source environment" and "Lab stock". We discuss the implications of these results within the discussion.

Table 4.9: Mixed effect models for each dataset for an effect of environmental lineage within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.13. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Data set	**Model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**% of Var.** σ^2_{ArtID}	**eff. samp.**	**pMCMC**
● cont. disp.	z ~Env. lin. - 1	Comm. Gard.	73.338	-0.178	(-0.836, 0.442)	0	1400	0.600
		Lab stock		-0.003	(-0.319,0.320)		1237	0.991
		Envir. source		-0.064	(-0.396, 0.292)		1400	0.677
● bin. all	g ~Env. lin. - 1	Lab stock	442.143	-0.244	(-0.680, 0.185)	0	1800	0.283
		Envir. source		-1.180	(-1.709, 0.715)		1800	$< 6\text{x}10^{-4}$ ***
● bin. all	g ~Env. lin. + random(Article ID) - 1	Lab stock	219.785	-0.204	(-0.670, 0.238)	80.54	1800	0.391
		Envir. source		-0.527	(-1.032, 0.024)		1800	0.049 *
● bin. - plants	g ~Env. lin. - 1	Lab stock	196.533	-0.235	(-0.390, -0.083)	0	1637	0.001 **
		Envir. source		-0.194	(-0.368, 0.006)		1800	0.045 *
● bin. - plants	g ~Env. lin. + random(Article ID) - 1	Lab stock	59.837	-0.187	(-0.458, 0.065)	78.18	1709	0.151
		Envir. source		-0.195	(-0.517, 0.087)		1800	0.198

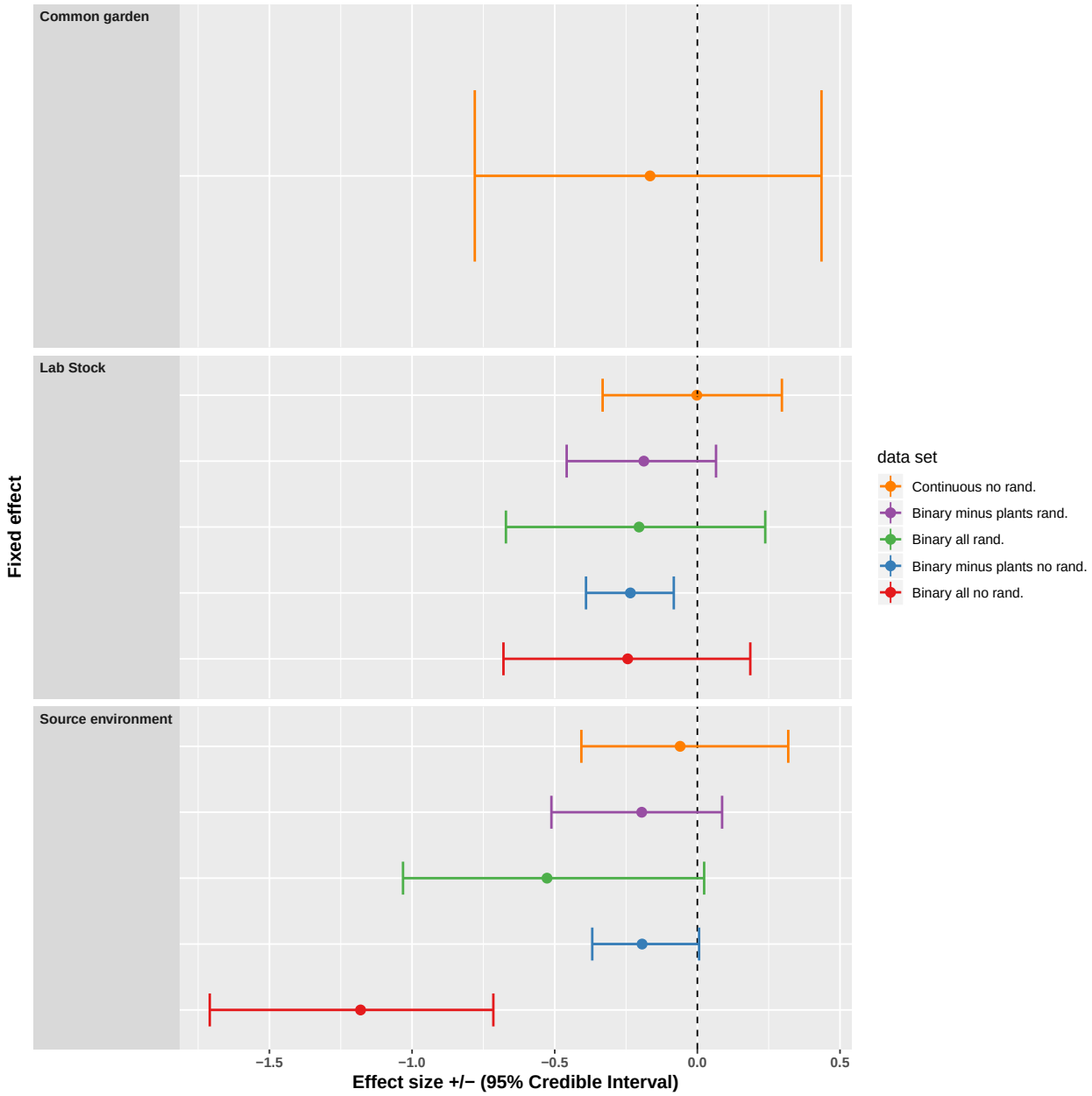

Figure 4.13: Forest plot of models fitted to all data sets with corresponding colour code given in table 4.9. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.6 Controlling for confounding variables

We have seen that there were significant results indicating that dispersal trade-offs exist within our data sets and that some of our explanatory variables could be responsible. However, recall from figures 4.1 and 4.2 that the proportions of factors within these explanatory variables are clustered around one category. This indicates a problem for any inference from the results described in section 4.5 since any one of our explanatory variables could, in fact, be a proxy for another. If this was the case, any results which were significant for both may in actuality be one effect. To illustrate this, within figure 4.2 we see that 67% of the species class is made up of Insecta and 39% of the dispersal investment is physical, which is potentially confounding since many insects invest in winged apparatus to disperse. The significant result we see in both of these effect sizes could be only one effect if Insecta is a proxy for physical investment, therefore, we would like to control for this possibility.

To control for these factors we used the binary minus plants data set, recall that article 1858 was considered an outlier. We then selected to use the lowest DIC of the models fitted to binary minus plants, which was the model that included article ID as a random effect. We then used this model exactly as described in section 4.4 with the caveat that different models for each explanatory variable had each of the remaining centred variables fitted an an additional fixed effect. By adopting this approach we could then investigate if the dispersal trade-offs seen in section 4.5 were still present when other potentially confounding factors were controlled for.

We only analysed those explanatory variables that had given significant results from the previous model, which are given below

- Dispersal type.
- Fitness type.

- Sex.
- Species class.
- Climate class.

We were unable to fit models for sex while controlling for species class and *vice versa* because the model could not estimate these as fixed effects. This is a common problem when variables are highly correlated. Therefore we do not consider this case any further.

Since all of the data sets in this analysis are identical, it is appropriate to assess each model based on the DIC. If a model with more parameters fitted provides a higher DIC than the original model – specified in the upcoming tables as *no control* – then we can reject it. We should reject the model since fitting more parameters loses explanatory power, and if it fits the data worse than the original model, then it explains less about the data with less information.

We calculated the proportional variance components for each model fitted to the data set, but do not report the proportions since they were similar to those of the models fitted to the same data in section 4.5.

4.6.1 Effect of dispersal type controlled

The best fitting model to the data was the fitness control model with a DIC of 46.983, while the original model had DIC of 47.695. The difference in DIC is less than 1, thus it would be sensible not to reject the non-control model. All models other than the fitness control had a higher DIC value and can be rejected.

Facultative dispersal

All models were estimated to have positive posterior mean coefficients with non-significant CIs (fig. 4.14, table 4.10). The effect of controlling for variables

widened the CIs and increased the effect size when controlling for fitness and climate. Therefore, there is still no evidence for a trade-off between facultative dispersal and non-dispersal traits.

Physical investment in dispersal

All models estimated negative posterior mean coefficients and significant CIs when controlling for sex and fitness, but non-significant when controlling for species class. The control for climate provided weakly significantly CIs. When controlling for other variables the effect was weakened, except for the sex control where the CIs were widened. Therefore, there is still evidence for a trade-off between investment in physical dispersal and non-fitness traits.

Table 4.10: Mixed effect models examining dispersal type while controlling for other factors within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.14. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Control variable	**Random effects (Art. ID) model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**eff. samp.**	**pMCMC**
no control	g ~disp. type - 1	Facultative	47.695	0.065	(-0.185, 0.320)	1800	0.616
		Phys. Invest.		-0.543	(-0.817, -0.267)	1800	$< 6\text{x}10^{-4}$ * **
● Species class	g ~disp. type + spec. class ctrl. - 1	Facultative	49.954	0.061	(-0.212, 0.378)	1800	0.667
		Phys. invest.		-0.435	(-0.960, 0.051)	1800	0.09 .
● Sex	g ~disp. type + sex ctrl. - 1	Facultative	55.895	0.050	(-0.264, 0.351)	1666	0.746
		Phys. invest.		-0.555	(-0.860, -0.237)	1800	0.001 **
● Fitness	g ~disp. type + fit. ctrl. - 1	Facultative	46.983	0.131	(-0.141, 0.401)	2231	0.356
		Phys. invest.		-0.484	(-0.757, -0.204)	1800	0.001**
● Climate	g ~disp. type + clim. ctrl. - 1	Facultative	59.267	0.215	(-0.058, 0.451)	1953	0.096 .
		Phys. invest.		-0.350	(-0.727, -0.002)	1625	0.056 .

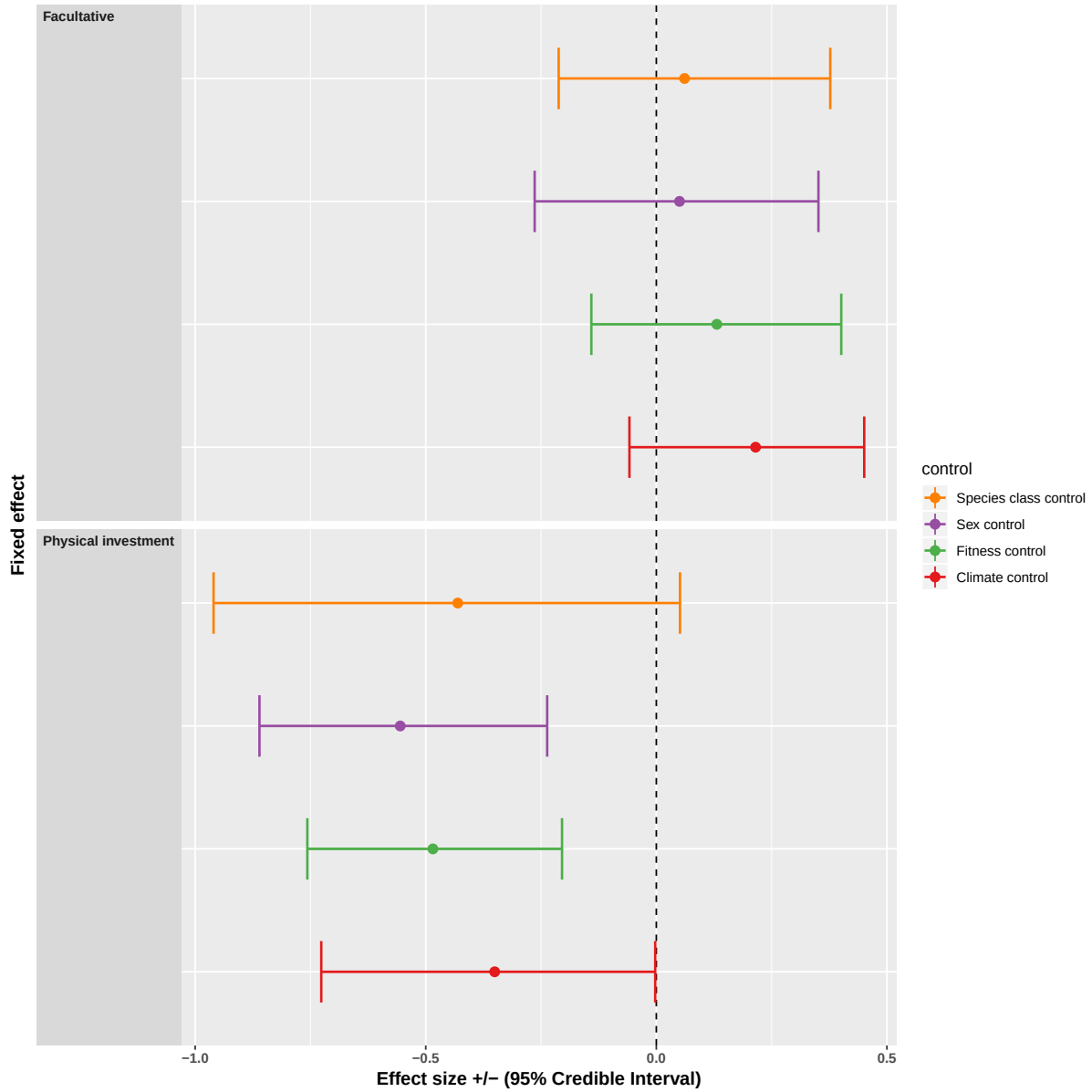

Figure 4.14: Forest plot of models fitted to the binary - plants data set with corresponding colour code given in table 4.10. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.6.2 Effect of fitness type controlled

The best fitting model to the data was the dispersal type control model with a DIC of 48.706, while the original model had a DIC of 56.055. Controlling for species class had a lower DIC (54.001)than the non-control model, indicating that species class helped explain the dispersal trade-off against the different fitness types.

Competitive

The models that controlled for sex and dispersal type were estimated to have negative posterior mean coefficients, whereas when controlling for species class and climate estimated positive coefficients (fig. 4.15 and table 4.11). All CIs were non-significant. The effect of controlling for sex and dispersal type weakened the effect, and in all cases widened the CIs. Therefore, when controlling for other variables, there is no evidence for a dispersal trade-off with competitive effects.

Fecundity

In all models the posterior mean coefficient was negative, except when controlling for climate (fig. 4.15 and table 4.11). Controlling for both species class and climate was non-significant, whereas for sex and dispersal estimates were significant. In all cases the effect was weakened and the CIs were widened. Therefore, there still remains evidence of a dispersal trade-off with fecundity.

Longevity

The models that controlled for sex and dispersal type were estimated to have negative posterior mean coefficients, whereas when controlling for species class and climate estimated positive coefficients (fig. 4.15 and table 4.11). None of the CIs were significant, but were widened when controlling for other variables. Therefore, there is still no evidence of a dispersal trade-off with longevity.

Table 4.11: Mixed effect models examining fitness type while controlling for other factors within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.15. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Control variable	**Random effects (Art. ID) model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**eff. samp.**	**pMCMC**
no control	g ~fitness - 1	Competitive	56.055	-0.120	(-0.467, 0.172)	1800	0.450
		Fecundity		-0.265	(-0.470, -0.034)	1611	0.021*
		Longevity		-0.095	(-0.384, 0.243)	1800	0.574
● Species class	g ~fitness + spec. class ctrl. - 1	Competitive	54.001	0.059	(-0.297, 0.454)	1800	0.762
		Fecundity		-0.050	(-0.336, 0.262)	1800	0.743
		Longevity		0.147	(-0.250, 0.542)	1938	0.459
● Sex	g ~fitness + sex ctrl. - 1	Competitive	59.003	-0.171	(-0.531, 0.162)	1800	0.313
		Fecundity		-0.372	(-0.669, -0.055)	1933	0.021*
		Longevity		-0.210	(-0.574, 0.198)	1800	0.289
●Dispersal	g ~fitness + disp. type ctrl. - 1	Competitive	48.706	-0.135	(-0.423, 0.161)	1800	0.355
		Fecundity		-0.292	(-0.474, -0.082)	1800	0.007 **
		Longevity		-0.109	(-0.406, 0.196)	1800	0.457
● Climate	g ~fitness + clim. ctrl. - 1	Competitive	67.872	0.142	(-0.199, 0.483)	1800	0.424
		Fecundity		0.035	(-0.242, 0.305)	1800	0.801
		Longevity		0.236	(-0.125, 0.597)	1800	0.198

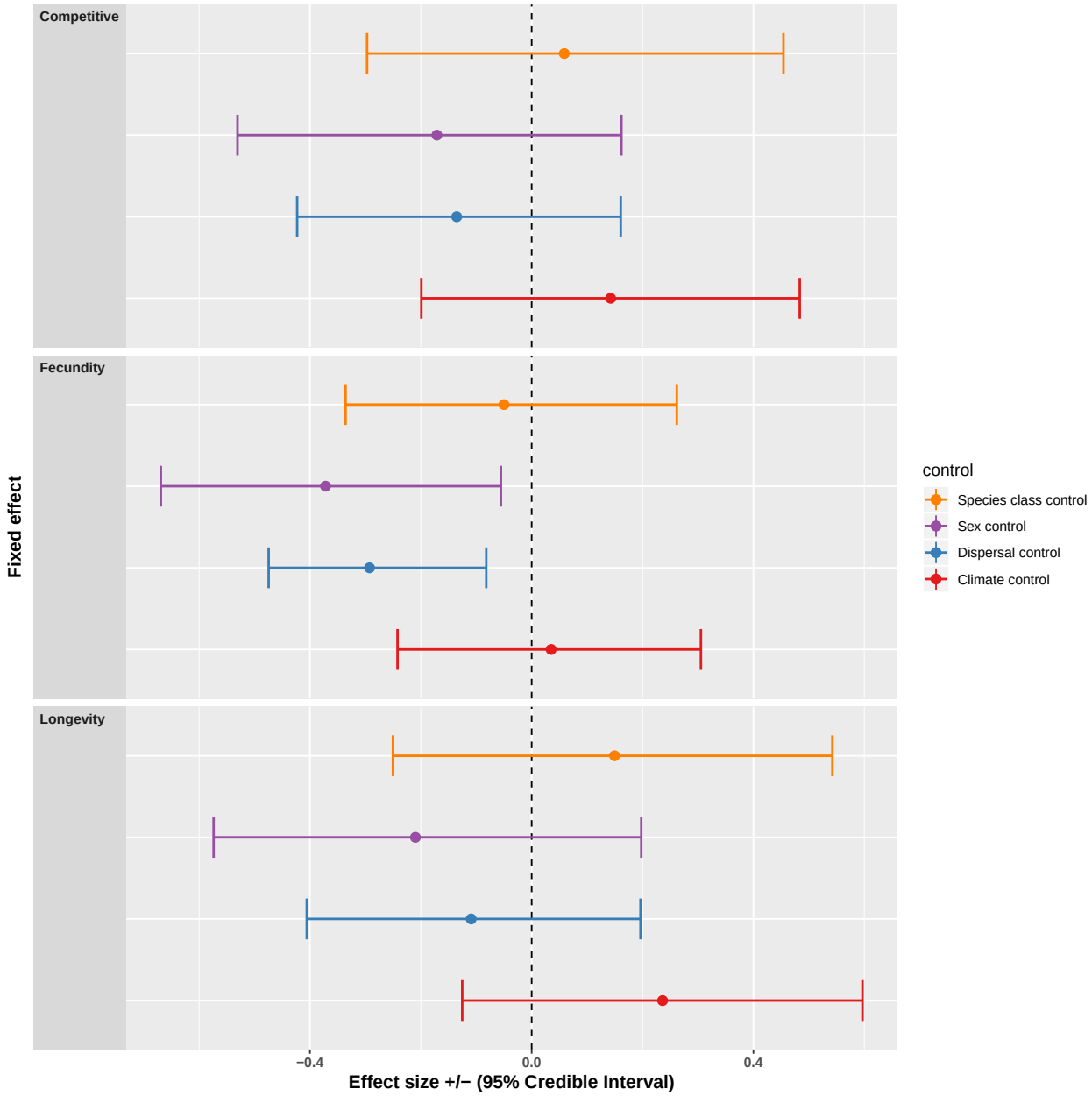

Figure 4.15: Forest plot of models fitted to the binary - plants data set with corresponding colour code given in table 4.11. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.6.3 Effect of sex controlled

The best fitting model to the data was the dispersal type control model with a DIC of 55.96, while the original model had a DIC of 58.487. Controlling for fitness did not noticeably improve the model fit, while controlling for climate type had a higher DIC and should be rejected.

Female

The models that controlled for dispersal type and fitness were estimated to have negative posterior mean coefficients, whereas when controlling for climate estimated positive coefficients (fig. 4.16 and table 4.12). Only when controlling for dispersal type were the CIs significant where they were also narrower, the remainder of the CIs were wider. Therefore, there is only evidence of a dispersal trade-off within females when controlling for dispersal type.

Male

The models that controlled for dispersal type and fitness were estimated to have negative posterior mean coefficients, whereas when controlling for climate estimated positive coefficients (fig. 4.16 and table 4.12). Only when controlling for dispersal type were the CIs significant where they were also narrower, the remainder of the CIs were wider. Therefore, there is no evidence of a dispersal trade-off within males when controlling for other factors.

4.6.4 Unknown

All models were estimated to have negative posterior mean coefficients with non-significant CIs (fig. 4.16 and table 4.12). Controlling for both dispersal and fitness weakened the effect and in all cases the CIs were wider. Therefore, there is no evidence of a dispersal trade-off within subjects with unknown sex.

Table 4.12: Mixed effect models examining sex while controlling for other factors within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.16. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Control variable	**Random effects (Art. ID) model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**eff. samp.**	**pMCMC**
no control	g ~Sex - 1	Female	58.487	-0.213	(-0.452, 0.009)	1800	0.063 .
		Male		-0.198	(-0.452, 0.069)	2028	0.148
		Unknown		-0.411	(-1.070, 0.277)	1800	0.218
● Fitness	g ~Sex + fit. ctrl. - 1	Female	58.474	-0.136	(-0.371, 0.141)	1800	0.287
		Male		-0.128	(-0.407, 0.153)	2112	0.382
		Unknown		-0.447	(-1.109, 0.275)	2228	0.210
● Dispersal	g ~Sex + disp. type ctrl. - 1	Female	55.960	-0.248	(-0.462, -0.024)	1800	0.034*
		Male		-0.221	(-0.435, 0.047)	1594	0.083 .
		Unknown		-0.305	(-0.927, 0.321)	1800	0.329
● Climate	g ~Sex + clim. ctrl. - 1	Female	68.401	0.153	(-0.152, 0.424)	1800	0.300
		Male		0.061	(-0.255,0.365)	1800	0.698
		Unknown		-0.228	(-0.902, 0.422)	1800	0.459

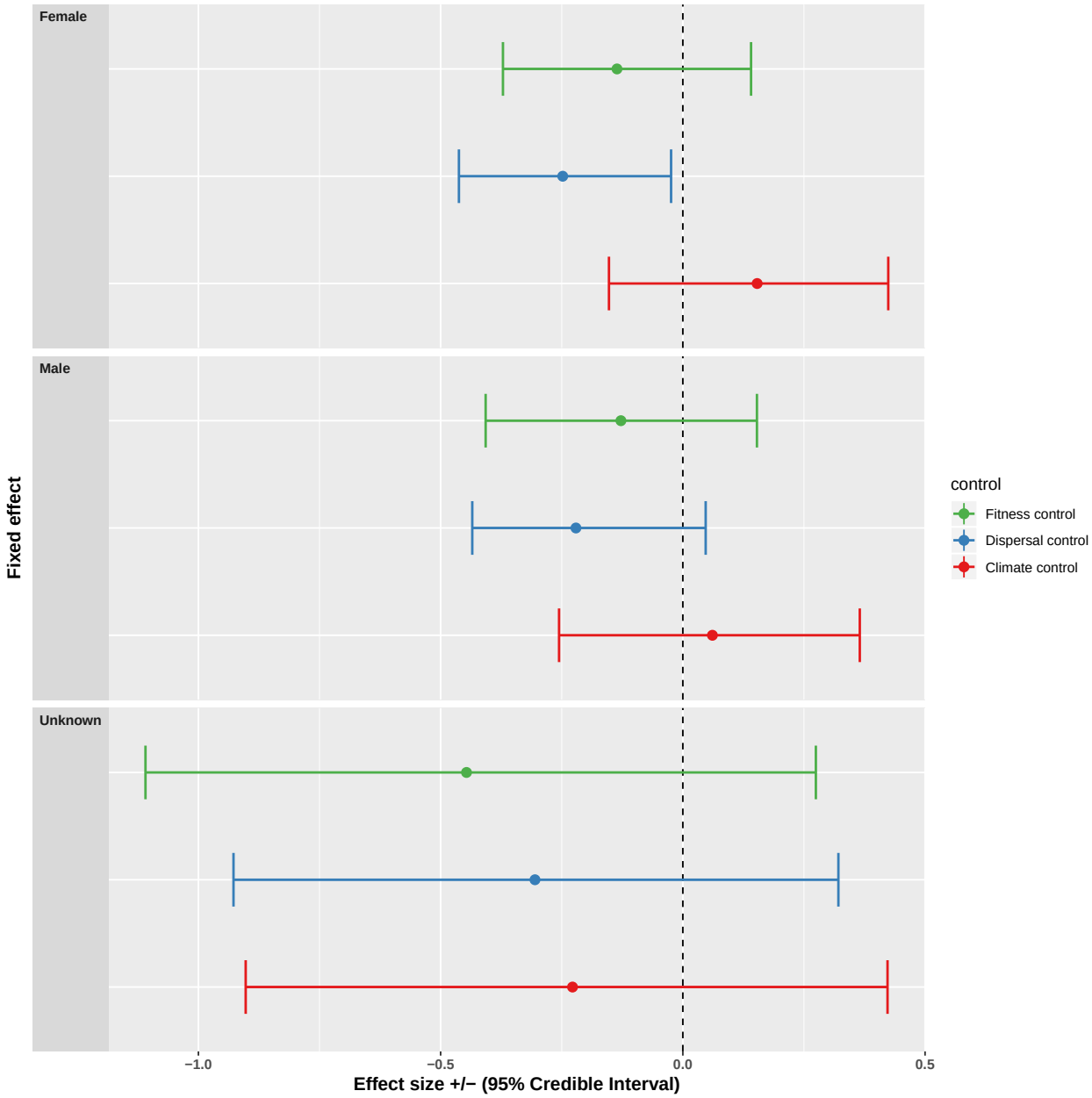

Figure 4.16: Forest plot of models fitted to the binary - plants data set with corresponding colour code given in table 4.12. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.6.5 Effect of species class controlled

The best fitting model to the data was the dispersal type control model with a DIC of 49.59, while the original model had a DIC of 51.862. The remainder of the models had a worse fit than the no control model and should be rejected.

Arachnida

The models that controlled for dispersal type and fitness were estimated to have negative posterior mean coefficients, whereas when controlling for climate type estimated positive coefficients (fig. 4.17 and table 4.7). None of the CIs were significant and all were widened. Therefore, there is no evidence of a dispersal trade-off within Arachnida when controlling for other variables.

Aves

The models that controlled for dispersal type were estimated to have negative posterior mean coefficients, whereas when controlling for fitness and climate were estimated as positive coefficients (fig. 4.17 and table 4.7). None of the CIs were significant and all were widened. Controlling for dispersal type strengthened the effect. Therefore, there is no evidence of a dispersal trade-off within Aves when controlling for other variables.

Insecta

All models were estimated to have negative posterior mean coefficients with significant CIs when controlling for fitness and dispersal type, but non-significant when controlling for climate type (fig. 4.17 and table 4.13). Controlling for fitness weakened the effect, but was strengthened when controlling for dispersal. All CIs were widened when controlling for other variables. Therefore, there is some evidence of a dispersal trade-off within Insecta when controlling for fitness and dispersal type.

Mammalia

All models were estimated to have positive posterior mean coefficients with weakly significant CIs when controlling for fitness, but non-significant when controlling

for climate type (fig. 4.17 and table 4.13). All CIs were widened when controlling for other variables. Therefore, there is no strong evidence of a dispersal trade-off within Mammalia when controlling for other variables.

Table 4.13: Mixed effect models examining species class while controlling for other factors within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.17. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are signif. codes: < 0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Control variable	**Random effects (Art. ID) model**	**Parameter**	**DIC**	**Estimated mean**	**95% CI intervals**	**eff. samp.**	**pMCMC**
no control	*g* ~Spec. class - 1	Arachnida	51.862	-0.353	(-1.162, 0.396)	1961	0.357
		Aves		-0.009	(-0.502, 0.493)	1800	0.974
		Insecta		-0.403	(-0.622, -0.154)	1800	0.002 **
		Mammalia		0.498	(-0.056, 1.073)	1800	0.086 .
● Fitness	*g* ~spec. class + fit. ctrl. - 1	Arachnida	53.766	-0.043	(-1.001, 0.754)	1800	0.927
		Aves		0.065	(-0.4458, 0.527)	2137	0.788
		Insecta		-0.334	(-0.573, -0.086)	1800	0.008 ***
		Mammalia		0.562	(-0.002, 1.128)	2137	0.051 .
● Dispersal	*g* ~spec. class + disp. type ctrl. - 1	Arachnida	49.590	-0.421	(-1.342, 0.430)	1240	0.351
		Aves		-0.259	(-0.736, 0.313)	1800	0.330
		Insecta		-0.301	(-0.542, -0.055)	1932	0.017 *
		Mammalia		0.240	(-0.339, 0.809)	1800	0.396
● Climate	*g* ~spec. class + clim. ctrl. - 1	Arachnida	63.525	0.179	(-0.618, 1.141)	1800	0.663
		Aves		0.316	(-0.374, 0.920)	2354	0.334
		Insecta		-0.119	(-0.489, 0.244)	1800	0.488
		Mammalia		0.269	(-0.408, 0.843)	1800	0.387

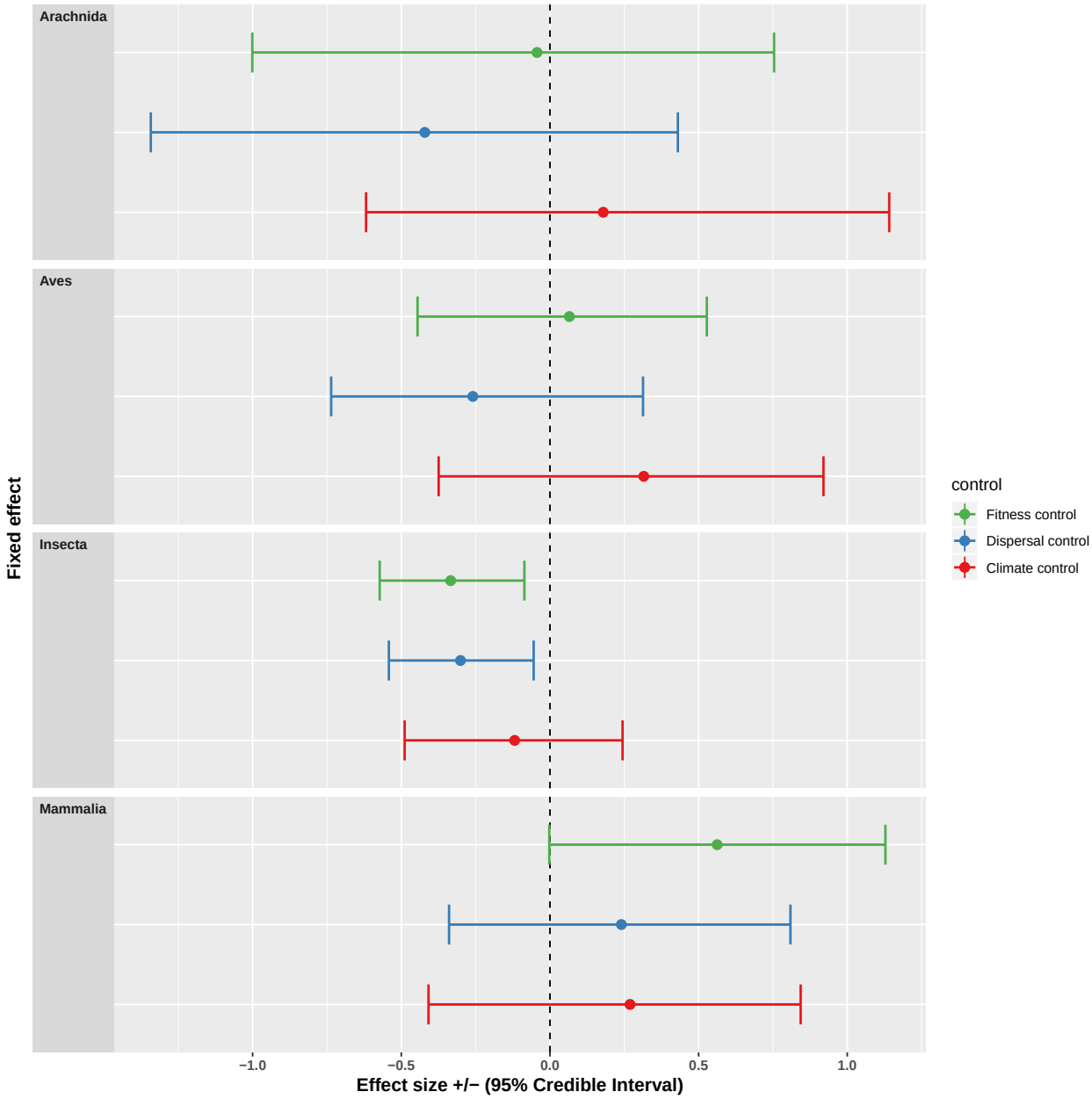

Figure 4.17: Forest plot of models fitted to the binary - plants data set with corresponding colour code given in table 4.13. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.6.6 Effect of Köppen-Geiger climate class controlled

The best fitting model to the data was the dispersal type control model with a DIC of 59.176, while the original model had a DIC of 65.245. Controlling for species class had a lower DIC than the no control model, suggesting that species class accounted for the variation within trade-offs explained by climate classification. The remainder of the models had higher DICs than the no control model and should be rejected.

Tropical

All models were estimated to have positive posterior mean coefficients with non-significant CIs, except when controlling for fitness which were significant (fig. 4.18 and table 4.14). When controlling for dispersal type and sex the effect was weaker, whereas for species class and fitness it was stronger. All CIs were widened when controlling for other variables, except for dispersal type where they were narrower. Therefore, there is no evidence of a dispersal trade-off within subjects from a tropical climate when controlling for other variables.

Dry

All models were estimated to have positive posterior mean coefficients with non-significant CIs (fig. 4.18 and table 4.14). When controlling for dispersal type the effect was weaker, whereas for the remainder it was stronger. All CIs were widened when controlling for species class, but were narrower for the remainder of the models. Therefore, there is no evidence of a dispersal trade-off within subjects from a dry climate when controlling for other variables.

Temperate

All models were estimated to have negative posterior mean coefficients with significant CIs except when controlling for species class (fig. 4.18 and table 4.14). When controlling for sex the effect was stronger, whereas for the remainder it was weaker. All CIs were narrower, except when controlling for sex. Therefore, there is evidence of a dispersal trade-off within subjects from a temperate climate, except when controlling for species class.

Continental

All models were estimated to have negative posterior mean coefficients with non-significant CIs (fig. 4.18 and table 4.14). When controlling for species class the effect was weaker, whereas for the remainder it was stronger. All CIs were wider than the non-control model. Therefore, there is no evidence of a dispersal trade-off within subjects from a continental climate when controlling for other variables.

Polar

All models were estimated to have negative posterior mean coefficients with non-significant CIs (fig. 4.18 and table 4.14). When controlling for species class and fitness type the effect was weaker, whereas for dispersal type and sex it was stronger. All CIs were wider for all models, except when controlling for dispersal type when they were narrower. Therefore, there is no evidence of a dispersal trade-off within subjects from a polar climate when controlling for other variables.

Table 4.14: Mixed effect models examining climate type while controlling for other factors within a dispersal trade-off. Models are for each data set with corresponding colour code with fig 4.18. Estimated mean from the posterior distribution and 95% credible intervals are given. Heterogeneity is given as a proportion of variance for both residuals and for random effects. Effective sample sizes from which the posterior means were sampled from. pMCMC values note the significance of CI intervals that do not significantly span 0. pMCMC significance codes are Signif. codes: pMCMC significance codes are Signif. codes: ¡0.001 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' '.

Control variable	Random effects (Art. ID) model	Parameter	DIC	Estimated mean	95% CI intervals	eff. samp.	pMCMC
no control	g ~climate - 1	Tropical	65.245	0.781	(-0.003, 1.606)	1800	0.058.
		Dry		0.4667	(-0.242, 1.133)	1800	0.172
		Temperate		-0.406	(-0.652, -0.179)	1800	0.001**
		Continental		-0.151	(-0.569, 0.2456)	1800	0.494
		Polar		-0.297	(-0.943, 0.387)	1800	0.368
● Species class	g ~climate + spec. class ctrl. - 1	Tropical	64.245	0.796	(-0.121, 1.660)	2058	0.083 .
		Dry		0.525	(-0.349, 1.291)	1800	0.201
		Temperate		-0.201	(-0.565, 0.132)	1800	0.251
		Continental		-0.164	(-0.677, 0.425)	1800	0.548
		Polar		-0.174	(-0.915, 0.522)	1800	0.6122
● Sex	g ~climate + sex ctrl. - 1 - 1	Tropical	68.187	0.672	(-0.167, 1.443)	1800	0.111
		Dry		0.545	(-0.137, 1.226)	1800	0.117
		Temperate		-0.528	(-0.826, -0.217)	1672	0.001*
		Continental		-0.273	(-0.715, 0.214)	1800	0.239
		Polar		-0.426	(-1.114, 0.288)	1800	0.228
● Fitness	g ~climate + fit. ctrl. - 1	Tropical	68.790	0.871	(0.101, 1.696)	1800	0.037 *
		Dry		0.500	(-0.203, 1.208)	1676	0.162
		Temperate		-0.355	(-0.596, -0.087)	1968	0.009**
		Continental		-0.080	(-0.502, 0.361)	1800	0.723
		Polar		-0.272	(-0.952, 0.365)	1800	0.411
● Dispersal	g ~climate + disp. type ctrl. - 1	Tropical	59.176	0.471	(-0.320, 1.150)	1800	0.192
		Dry		0.422	(-0.265, 1.019)	1551	0.187
		Temperate		-0.367	(-0.590, -0.152)	1800	$< 6\text{x}10^{-4}$ ***
		Continental		-0.269	(-1.161, 0.103)	1800	0.163
		Polar		-0.577	(-1.161, 0.103)	1800	0.078 .

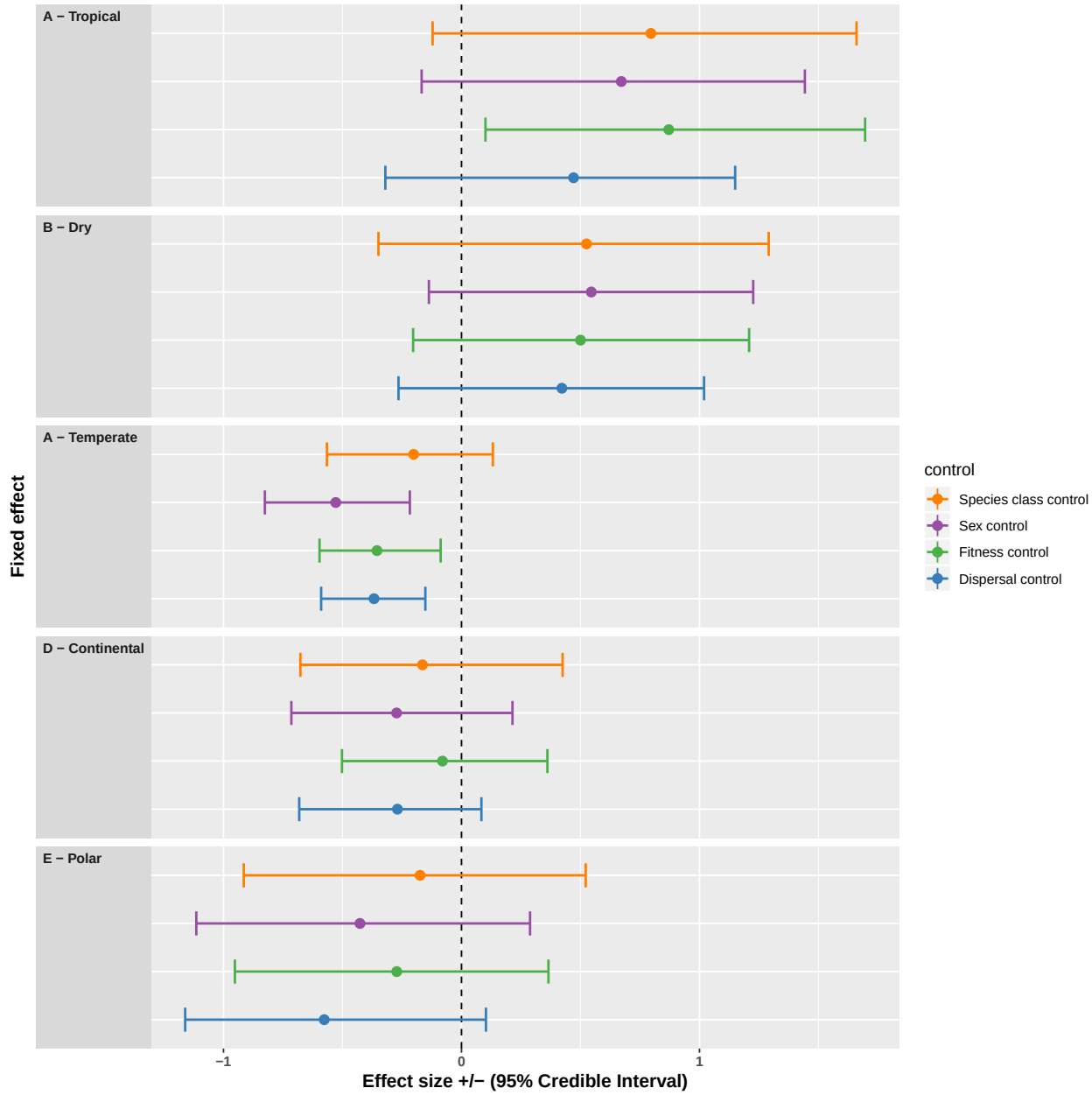

Figure 4.18: Forest plot of models fitted to the binary - plants data set with corresponding colour code given in table 4.14. Effect sizes are given with 95% credible interval bars. A line is included indicating 0 where any point with a negative effect size is considered a trade-off. Each level of the explanatory variable is plotted separately with results for estimates from each parameter included within its respective sub-plot.

4.7 Discussion

Throughout the course of this chapter we have shown that dispersal trade-offs exist within the empirical literature. We have aggregated the available empirical measurements that were present in the literature into a database for other researchers to utilise, and hopefully add to in the future. The inclusion criteria we have developed can also be used by other researchers to repeat our analysis, or to build on our findings with their own adaptations. From the data set we built from the literature, we now have a picture of what the proportions of studies examined measured (figs. 4.1 & 4.2), which we can use to determine where the gaps in the literature are for future research. Using the data set, we were able to establish that dispersal trade-offs were present and which factors best explained the results, furthermore, these were still significant when controlling for potentially confounding factors.

Through our analysis we have established that dispersal trade-offs were present in all of our data sets, except within the continuous dispersal data. When we accounted for the heterogeneity of the effect sizes using our fixed effects, we found that investments in

- physical investment in dispersal,
- fecundity,
- females,
- insects,
- plants,
- temperate climates,

significantly explained the trade-off. We also found that accounting for variation between articles as random effects provided better fitting models, but weakened the trade-off effect. The variation between articles also accounted for the majority of variation within the effect sizes – typically 70% or higher.

We note a special case within the taxonomic group Mammilia of a significantly positive effect size. When an effect size is positive it implies that, rather than a trade-off, we have the opposite effect where an increase in dispersal is accompanied by an increase in non-dispersal traits. A positive effect then implies that mammals that are more dispersive also have higher fitness traits.

Since the data within our explanatory variables were clustered, any models that provided significant results were then fitted with additional fixed effects to control for correlation between explanatory variables. We found that the general trend of the models were to reduce the trade-off and widen any CIs, with few exceptions. We discovered significant trade-offs in

- physical investment in dispersal
  - when controlling for fitness, sex, and climate,
- fecundity
  - when controlling for dispersal type and sex,
- females
  - when controlling for dispersal type,
- Insecta

  - when controlling for dispersal type and fitness,

- temperate climates

  - when controlling for dispersal type, fitness, and sex.

The best fitting model was dispersal type controlling for fitness type – DIC = 46.983. We found that those models controlling for dispersal type consistently provided a better fit to the data – lowest DIC – suggesting that dispersal type was a significant factor in determining a dispersal trade-off. We found that controlling for species class and climate always produced non-significant results, with the exception of controlling for climate within dispersal type. This suggests that species class and climate are correlated with the other factors. We can therefore conclude that dispersal trade-offs in the aforementioned proxies exist, and are particularly strong when physically investing in dispersal apparatus.

With reference to Mammalia, there was a positive significant result when controlling for fitness type, but non-significant for climate and dispersal type. This result suggests that when controlling for fitness there seems to be an increase in both dispersal and non-dispersal traits, but since we have seen that dispersal type is the best predictor, and moreover that when controlling for dispersal, the trade-off became non-significant.

When we examined the environmental lineage of the subjects within the binary minus plants data set with random effects included, we found that there was no significant difference betwee laboratory stock and subjects sourced from their natural environment. This is an significant finding because it implies that the dispersal trade-off within an organism is unaltered after generations of lab conditions which are unlike its natural environment. Our findings suggest that results examining dispersal trade-offs using laboratory specimens are robust to scrutiny of their rearing conditions influencing trade-off results of such experiments.

Our results, although significant, come with caveats inherent in the structure of the data. Much of the levels of variation within subjects of our data – our explanatory variables – were clustered around one factor, e.g. Insecta within species class, or fecundity within fitness proxy, which could imply that that the trade-offs we seen are a consequence of the data itself, rather than true effects, despite controlling for other potentially confounding effects. Ideally, it would have been beneficial to have less aggregation within factors, but this is the current state of research and our results should be seen within this context.

In this chapter we have established the existence of general dispersal trade-offs in a quantified manner, which the literature is currently missing. This study gives weight to the conjecture that dispersal trade-offs exist and should be utilised within life-history theory as a biological constraint. The study of range expansions and invasions are difficult and costly to investigate empirically, thus theoretical models are currently the only viable method to understand these complex events. We believe that accounting for a dispersal trade-off will provide theoretical predictions for range expansions, for which dispersal related life histories is critical, with more representable estimates of natural systems.

Within the context of this thesis, our theories predicted faster invasions when organisms were polymorphic and subject to a dispersal-reproduction trade-off. This chapter established that fecundity partly explained a dispersal trade-off (fig. 4.9, 4.15, & tables 4.5, 4.11), which provides weight to our assumption of a dispersal-reproduction trade-off. The implications are then that accounting for a trade-off could explain some of the variation within estimates of invasion speed – often under estimates – that are so widely reported within the literature [78].

Chapter 5

Anomalous invasion speeds of sexually reproducing species

5.1 Introduction

In chapters 2 & 3 we have concerned ourselves with haploid models where reproduction is clonal, and traits are inherited from the parent with a small chance of mutatiting between strains. We witnessed that anomalous invasion speeds were possible under certain conditions. However, in nature sexual reproduction covers a significant proportion of reproducing organisms, particularly those that are of interest to invasion management, and conservationists. Therefore, we would like to determine if changing the reproductive process affects the realisation of anomalous invasion speeds.

Previous research of spatial spread with sexually reproducing species is uncommon [32]. Those studies have assumed simplifying assumptions, such as limiting breeding between genotypes and excluding interaction terms between them [194]. The focus of other studies has been on tracking allele frequencies, or the spread of alleles within populations, but not the overall abundances of the organisms themselves, which is the primary interest within invasion modelling [95].

Models that do account for abundances exist, but do not account for differences in dispersal ability [119, 120].

Since we want to investigate how sexual reproduction affects anomalous invasion speeds, the assumptions we mentioned previously will not be sufficient for our purpose. We are interested in the case when genotypes within populations inter-breed and interact, e.g. through competition, which should occur within natural species. To investigate spatial invasions, or range expansions, it is necessary to account for how the species propagates through space. Solely examining the genetics of the population does not capture the spatial dynamics, only the traits which may give rise to them. Furthermore, the previous issues we have raised cannot address variation within species and the consequences of dispersal polymorphism during invasion.

The technical implications for studying sexual reproduction are that the interactions that give rise to births are intrinsically non-linear. This means that linearisation is incapable of providing tractable analytical solutions. There is a further issue when numerically simulating sexual reproduction related to non-linearity, which creates singularities within continuous space and time models. If any of the individuals disperse to where there is no mate, then the model will divide by 0 and the solution becomes undefined. Therefore, we use simulations within a discrete time and space framework rather than PDEs or analysis which we have used previously. Although the Lotka-Volterra framework we have used in chapters 2 & 3 is problematic for sexual reproduction, models that are discrete in both time and space have proven to be powerful tools for studying such systems since the dynamics of the discrete system are similar to the continuous case [36].

We seek to address this gap in the literature by providing a discrete time and space invasion model that accounts for sexual recombination, where genotypes experience variation in their trait expressions of growth and dispersal. To account for sexual recombination, we shall assume the organism is diploid and reproduces according to Mendelian inheritance with genotypes AA, BB, and AB [121]. We

shall examine a system where the homozygotes have either strong reproductive and weak dispersal abilities, and *vice versa*. This will then allow us to investigate the effect that the heterozygote has on invasion speeds when its trait expressions differ.

Using simulations we will ascertain if anomalous invasion speeds are still possible when the means of inheritance is through sexual recombination, rather than direct inheritance from the clonal parent. First we shall establish whether travelling wave solutions exist, since without their existence there can be no comparison with the previous studies. Then we shall determine if anomalous invasion speeds are realised. Finally, we shall determine what impact the heterozygote has on invasion when it: mimics either homozygote (dominance of one allele); or has a combination of the weakest, strongest, and average of dispersal and reproduction traits.

We shall also examine the effects of varying degrees of mutation between alleles, i.e. A mutates into B. Mutation was critical in our previous chapters because it is the only means of changing phenotype in a clonal reproductive system. Switching between phenotypes by sexual recombination does not require mutation, therefore, it is unclear if mutation in this system is as critical as we have previously seen.

5.2 Model formulation

To investigate sexual reproduction during invasions we shall develop a model that is discrete in both time and space. We shall assume that matings happen locally and at random where all genotypes are able to pair with each other without limitations, i.e. no assortative mating between genotypes. Sexual recombination will behave according to Mendelian inheritance where the homozygotes are denoted AA and BB, and the heterozygote as AB. We include a probability that a single gamete can mutate at birth from A into B and *vice versa*. We also assume that generations can over-lap, so that we can draw comparisons with chapters 2 & 3.

We shall define the model as a difference equation that describes each stage of a generation cycle chronologically as:

1. population growth and mutation dynamics,

2. density-dependent mortality,

3. dispersal.

5.2.1 Population growth and mutation

Let us begin by defining the population as

$$N^t_{i,x} = F^t_{i,x} + M^t_{i,x}, \tag{5.1}$$

where N is the total population, F and M are the population of females and males. Here $i \in S = \{AA, AB, BB\}$ denotes the genotype, x is the position in space, and t is the generation number. We can now define the reproductive step with respect to each gamete as

$$\begin{aligned} G^t_{fA,x} &= \big(r_{AA}F^t_{AA,x} + \frac{1}{2}r_{AB}F^t_{AB,x}\big), \qquad & G^t_{fB,x} &= \big(r_{BB}F^t_{BB,x} + \frac{1}{2}r_{AB}F^t_{AB,x}\big), \\ P^t_{mA,x} &= \frac{M^t_{AA,x} + \frac{1}{2}M^t_{AB,x}}{\sum_{i\in S} M^t_{i,x}}, & P^t_{mB,x} &= \frac{M^t_{BB,x} + \frac{1}{2}M^t_{AB,x}}{\sum_{i\in S} M^t_{i,x}}. \end{aligned} \tag{5.2}$$

Where $G^t_{fA,x}$ is the number of female produced gametes of allele A at position x and generation t, and similarly for allele B; $P^t_{mA,x}$ is the probability that a male produced gamete is allele A at position x at generation t, and similarly for B; r_i is the proportion of offspring produced per generation. We now introduce the possibility of mutation between alleles. To make this model comparative to those in our previous chapters we shall assume that a constant proportion of offspring mutate at birth from one allele to another, i.e. A mutates into B at a rate μ and

vice versa for B into A. Here we have assumed that the rate of mutation between alleles is equal. This is then represented as

$$\begin{aligned} \hat{G}^t_{fA,x} &= G^t_{fA,x}\big(1-\mu\big) + G^t_{fB,x}\mu, \qquad \hat{G}^t_{fB,x} = G^t_{fB,x}\big(1-\mu\big) + G^t_{fA,x}\mu, \\ \hat{P}^t_{mA,x} &= P^t_{mA,x}\big(1-\mu\big) + P^t_{mB,x}\mu, \qquad \hat{P}^t_{mA,x} = P^t_{mA,x}\big(1-\mu\big) + P^t_{mB,x}\mu, \end{aligned} \tag{5.3}$$

where $\hat{G}^t_{fA,x}$ is the number of female gametes of allele A at position x and generation t accounting for mutation, and similarly for allele B, and $\hat{P}^t_{mA,x}$ is the probability of encountering a male of allele A at position x at generation t accounting for mutation, and similarly for B. We now define the number of offspring for each genotype at generation t, $R^t_{i,x}$ for $i \in S$ as

$$\begin{aligned} R^t_{AA,x} &= \hat{G}^t_{fA,x}\hat{P}^t_{mA,x}, \\ R^t_{AB,x} &= \hat{G}^t_{fA,x}\hat{P}^t_{mB,x} + \hat{G}^t_{fB,x}\hat{P}^t_{mA,x}, \\ R^t_{BB,x} &= \hat{G}^t_{fB,x}\hat{P}^t_{mB,x}, \end{aligned} \tag{5.4}$$

which we then use to define growth per generation as

$$\tilde{F}^t_{i,x} = R^t_{i,x}\alpha + F^t_{i,x}, \quad \tilde{M}^t_{i,x} = R^t_{i,x}\big(1-\alpha\big) + M^t_{i,x}, \tag{5.5}$$

where $\tilde{F}^t_{i,x}$ is the new population of females after reproduction for genotype $i \in S$, and similarly for males $\tilde{M}^t_{i,x}$, and $\alpha \in [0,1]$ is a constant parameter which controls the sex ratio.

5.2.2 Density-dependence

We now turn our attention to density-dependent mortality. To account for this we shall use a version of the Beverton-Holt model that accounts for interacting species known as the Leslie-Gower model [104]. This model has been shown to be the discrete time and space model for the continuous competition dynamics we used within chapters 2 & 3 [103]. We shall now combine equn. (5.5) with the Leslie-Gower model to give our density dependence relation for both males and

females at generation t, which is given as

$$\begin{aligned}\hat{F}^t_{i,x} &= \frac{\tilde{F}^t_{i,x}}{1+\sum_{j\in S} C_{ij}\big(\tilde{F}^t_{j,x}+\tilde{M}^t_{j,x}\big)},\\ \hat{M}^t_{i,x} &= \frac{\tilde{M}^t_{i,x}}{1+\sum_{j\in S} C_{ij}\big(\tilde{F}^t_{j,x}+\tilde{M}^t_{j,x}\big)},\end{aligned} \tag{5.6}$$

where $i, j \in S$, and C_{ij} are the competitive effects between genotypes i and j.

5.2.3 Dispersal dynamics

Finally, we shall address the dispersal dynamics. Here we have assumed a constant proportion of individuals with genotype $i \in S$ equally disperse to neighbouring cells with some probability D_i, whereas the remainder linger within their current cell. The population of males and females at position x for the next generation $t+1$ is given as

$$\begin{aligned}F^{t+1}_{i,x} &= \frac{D_{i,f}}{2}\hat{F}^t_{i,x-1} + \frac{D_{i,f}}{2}\hat{F}^t_{i,x+1} - D_{i,f}\hat{F}_{i,x} + \hat{F}^t_{i,x},\\ M^{t+1}_{i,x} &= \frac{D_{i,m}}{2}\hat{M}^t_{i,x-1} + \frac{D_{i,m}}{2}\hat{M}^t_{i,x+1} - D_{i,m}\hat{M}_{i,x} + \hat{M}^t_{i,x},\end{aligned} \tag{5.7}$$

where $D_{i,f}$ is the proportion of females dispersing for genotypes $i \in S$, and similarly for males.

We have thus defined a general model in discrete time and space that accounts for the populations of males and females that allow different dispersal strategies for genotypes and sex.

5.3 Numerical experiments

Using the difference equation model defined in section 5.2, we can now investigate the effects of sexual recombination on invasion using numerical experiments. Although we have derived a general model, we shall now restrict ourselves to a special case. Consider a population as described in section 5.2 where the genotype AA has superior reproductive abilities, but inferior dispersal ability to the other homozygote BB. That is, there is a dispersal reproduction trade-off between geno-

types. This then provides us the opportunity to investigate the impact that the heterozygote AB has on invasion dynamics depending on what traits it adopts. We shall assume that there are no differences in sex, i.e. males and females within the same genotype do not differ in dispersal or competitive ability. Since fitness is defined solely by the females, males do not have a direct impact on fitness other than their relative abundance. We shall assume that the sex ratio is equal, which is to say that $\alpha = \frac{1}{2}$. By simplifying the model in this way, there is no distinction between males and females, therefore we shall consider the densities of the genotypes only, i.e. N_{AA}, N_{BB}, and N_{AB}.

We would like to draw comparisons between the dynamics in chapters 2 & 3, moreover, that if changing the reproductive system from clonal to involve sexual recombination retains the anomalous invasion speed phenomenon. We shall then assume the same environmental conditions from chapter 2 that the habitat is favourable to the species, and is unoccupied by a resident. Therefore, for each case, we must establish if travelling wave solutions exist, and if the speed of invasion of all genotypes travels at a single speed which is faster than either homozygote in isolation. In the previous chapters we have seen that mutation between strains was critical for a system to exhibit anomalous invasions speeds, recall that when there is no mutation that each strain invades at a speed specified by the F-KPP speed [47, 53, 99]. However, to change between strains in this system does not require mutation. To investigate if mutation is as critical as we have previously seen, we shall investigate different mutation values and compare the speeds of invasion for each case.

Numerical simulations were conducted in MATLAB 2016a [116] using the difference equation defined in section 5.2. The invasion speed was calculated by recording 100 time points over the total temporal period and noting the spatial coordinates of the the midpoint of the population density wavefront. The speed between each temporal point was calculated – speed = $\frac{\text{distance}}{\text{time}}$ – and an average of those speeds was taken. The general parameters and initial conditions of the experiments are given in table 5.1. The parameter values for the heterozygote

have not been defined in table 5.1 since we will be examining the consequences of which traits it adopts for each case.

When presenting the simulation results, we have normalised the carrying capacity of the species to be 1, the normalisation factor is calculated from the denominator of equn. (5.6) and was calculated prior to each separate simulation.

5.3.1 Monogamete case

To investigate the monogamete case, the reproduction and dispersal parameters given in table 5.1 were chosen such that each genotype in isolation would invade at the same speed. We can see from table 5.2 that this is the case to within 2%. If we examine panels c and d from fig. 5.1, we can see in c.1 and d.1 that both monogamete populations invade at the same speed. Panels c.2 and d.2 show the the plot of the travelling wave variable $z = x - ct$, where x is position in space, c is the calculated wave speed, and t the generation time. From here we can conclude that the solution exhibits travelling wave solutions.

5.3.2 Dominance of reproduction

Here we examine the case of dominance in the reproductive trait, i.e. the heterozygote has identical traits to the reproductive homozygote, AA. Since there is now more than one genotype, we introduce mutation between alleles as described in section 5.2. We can see from table 5.3 that the invasion speed when all genotypes are together is indeed faster than either homozygote in isolation. We see that mutation increases the average speed when the heterozygote experiences dominance within the reproductive allele A. We should also note that when mutation is 0, the invasion proceeds as a stacked front, which is when two invasions fronts are travelling at different speeds. If none of these speeds are faster than either homozygote in isolation then there is no anomalous invasion speed. In this case it is the heterozygote AB and homozygote AA that are travelling at the slower speed of $c_1 = 0.1448$ and the dispersive homozygote BB that is invading at speed $c_2 = 0.2197$ – it's speed in isolation. This case can be compared with the monogamete cases by examining panels b, c and d in fig. 5.1, where panel b

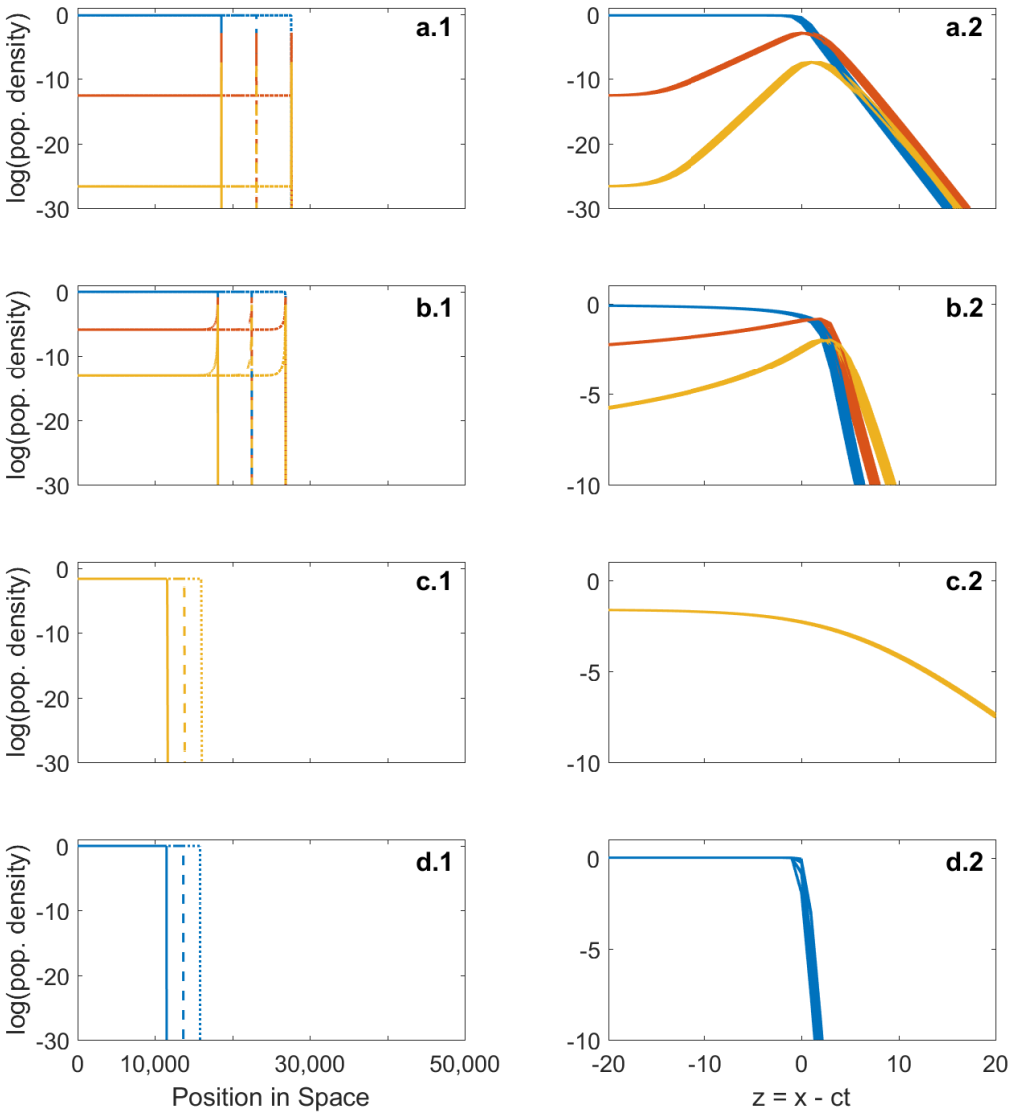

Figure 5.1: Invasion of vacant habitat by all genotypes. The natural logarithm of the population density is given on the y-axis. Position in space is given on the x-axis. The reproductive genotype is shown in blue, the dispersive genotype is shown in orange, and the heterzygote is shown in yellow. Parameter values and initial conditions are given in table 5.3 and $\mu = 1\text{x}10^{-6}$. Panels in row a show the invasion when the dispersive allele is dominant. Panels in row b shows when the reproductive allele is dominant. Panels in row c shows the monogamete case for the dispersive genotype. Panels in row d shows the monogamete case for the reproductive gamete. Panels in the first column depict invasion fronts at 3 equidistant time shots at $t = 30,000$ (solid line), $t = 40,000$ (dashed line), and $t = 50,000$ (dotted line). Panels in the second column depict the natural log. of the population density against the wave variable $z = x - ct$ of each case at 10 time points, 5,000 time steps apart.

displays the case when $\mu = 1\text{x}10^{-6}$. The case when $\mu = 0$ is shown in fig. 5.2.

5.3.3 Dominance of dispersal

We now examine the results of our numerical experiments when there is dominance of the dispersive trait, i.e. the heterozygote has identical traits to the dispersive homozygote, BB. From table 5.3 we can see that just as in the previous case, there is a single invasion speed that the population invades at, which is faster than either homozygote in isolation. Therefore, we have evidence of an anomalous invasion. We also notice that the rate of mutation between alleles has an almost indistinguishable influence on the observed invasion speeds. This case differs from the previous case in that there is no stacked front when mutation is 0. We can compare the case when there is dominance in the dispersive allele B with those of the monogamete cases in panels a, c, and d fig 5.1. We can also see from fig. 5.1 the existence of travelling wave solutions in panel a.2.

We shall not examine the extreme case when the reproductive homozygote is sessile, and the dispersive homozygote is infertile, and any dominance of either allele. Upon deeper reflection, we will come to the conclusion that no invasion will take place since any dispersive genotypes will produce no progeny, and those that do will be unable to disperse.

5.3.4 Strong heterozygote

We now turn our attention to the case when the heterozygote expresses the stronger traits of both homozygotes, i.e. the dispersive abilities of BB, and the reproductive abilities of AA. If we examine table 5.4 we find that the population travels at a single invasion speed which is faster than either homozygote in isolation, which should not surprise us since there is no trade-off. This case is analogous to that discussed in chapter 2.4.2. when the trade-off curve has negative curvature and the invasion should proceed at the speed of the optimum phenotype. We can also see from table 5.4 that mutation has very little impact on the invasion speed. The results of this experiment can be seen in panel a of fig. 5.3, where we can also see visual evidence of travelling wave solutions in panel

Table 5.1: General parameter values used within numerical simulations.

Descriptor	**Symbol**	**Values**
Total landscape size	L	[0,50000]
Total temporal period	T	50000
Proportion dispersing	(D_{AA}, D_{BB})	$(0.00055, 0.5)$
Per capita growth rate	(r_{AA}, r_{BB})	$(5, 0.1)$
Initial conditions	$N_i^* = [0, 5000]$	$\frac{1}{6}$

Table 5.2: Average invasion speeds of each monogamete case in isolation.

Genotype strain	**Average speed**
Reproductive strain AA	0.2164
Dispersive strain BB	0.2197

Table 5.3: Average invasion speeds at different rates of mutation when the heterozygote expresses dominance of either allele.

Heterozygote traits	**Mutation parameter**	**Average speed**
Dominance of reproduction AB = AA	$\mu = 0$	$c_1 = 0.1448$, $c_2 = 0.2197$
	$\mu = 1\text{x}10^{-6}$	0.4374
	$\mu = 0.001$	0.4431
	$\mu = 0.01$	0.4535
	$\mu = 0.1$	0.4847
Dominance of dispersal AB = BB	$\mu = 0$	0.4525
	$\mu = 1\text{x}10^{-6}$	0.4525
	$\mu = 0.001$	0.4524
	$\mu = 0.01$	0.4514
	$\mu = 0.1$	0.4447

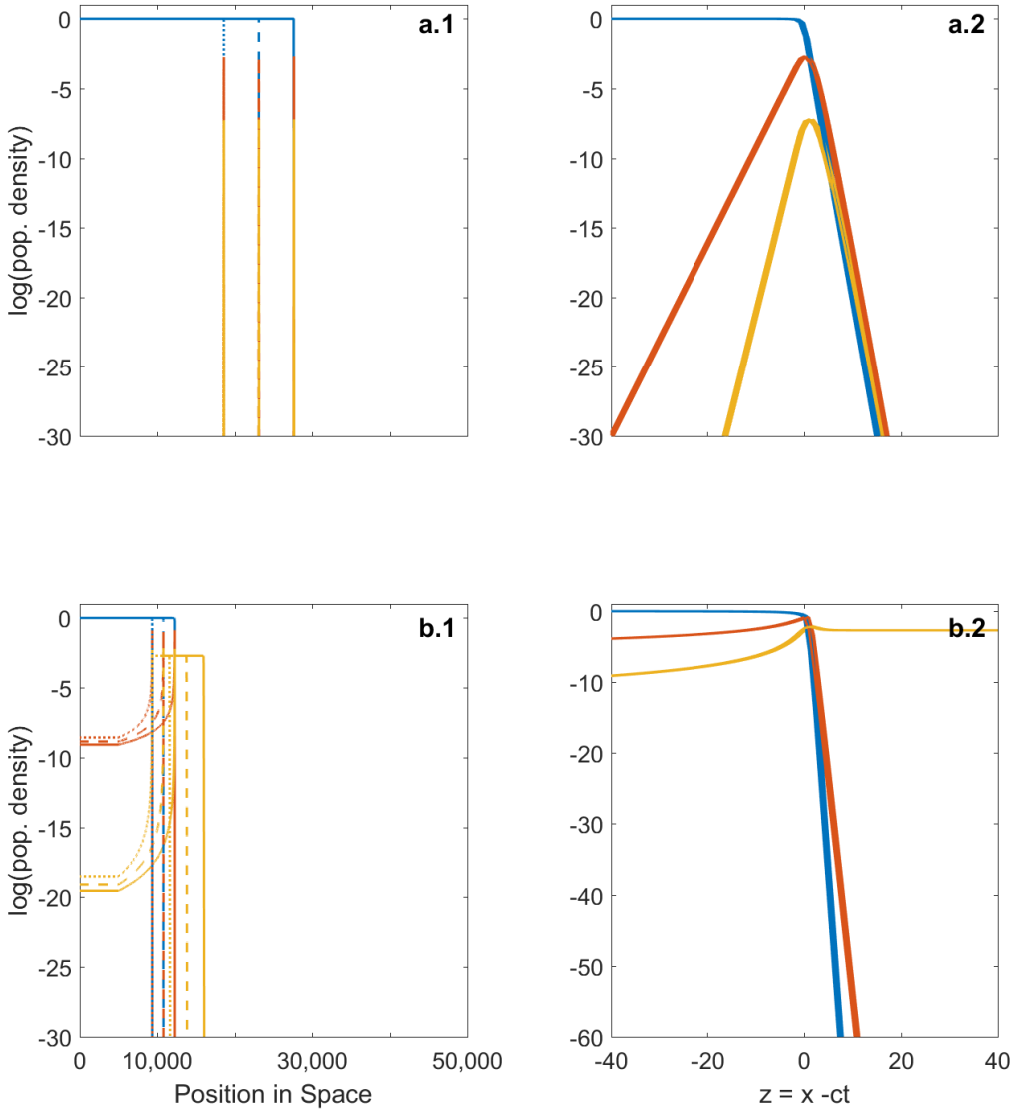

Figure 5.2: Invasion of vacant habitat by all genotypes. The natural log. of the population density is given on the y-axis. Position in space is given on the x-axis. The reproductive genotype is shown in blue, the dispersive genotype is shown in orange, and the heterzygote is shown in yellow. Parameter values and initial conditions are given in table 5.3 and $\mu = 0$. Panels row a show the invasion when the dispersive allele is dominant. Panels row b shows when the reproductive allele is dominant. Panels a.1 and b.1 depict invasion fronts at 3 equidistant time shots at $t = 30,000$ (solid line), $t = 40,000$ (dashed line), and $t = 50,000$ (dotted line). Panels a.2 and b.2 depict the natural log. of the population density against the wave variable $z = x - ct$ of each case at 10 time points, 5,000 time steps apart..

a.2. The strong heterozygote invades faster than both the intermediate and weak heterozygotes.

5.3.5 Intermediate heterozygote

We shall now examine the scenario when the heterozygote expresses traits at the midpoint between the reproductive and dispersal abilities of both homozygotes. If we review table 5.4, we will see that population invades at a single speed which is faster than that of either homozygote in isolation, which is evidence of an anomalous invasion speed. We should also notice that the mutation rate has little influence on the invasion speed, with marginal increases as the mutation rate increases in order of magnitude. The results of this experiment can be seen in panel b of fig. 5.3, where there is visual evidence of travelling wave solutions in panel b.2. The invasion speed of the intermediate heterozygote was slower than that of the strong, but faster than that of the weak heterozygote.

5.3.6 Weak heterozygote

Finally, we address the case when the heterozygote expresses the weaker traits of both homozygotes, i.e. the dispersive abilities of AA, and the reproductive abilities of BB. Upon inspection of table 5.4 we can see that the population travels at two distinct speeds when mutation is low enough, indicating that there exists a stacked front. Moreover, the slower wave comprised of the reproductive homozygote, and the faster wave of the heterozygote increased in speed with increasing mutation. However, when mutation passes a threshold – $\mu = 0.001$ – we again see the population invading at a single speed. Just as we have seen in the case when the reproductive allele is dominant – section 5.3.2 – there is no anomalous invasion speed when mutation is too low. In fact, we only see an anomalous invasion speed when mutation is relatively high, $\mu = \mathcal{O}(0.1)$. Visual results of this experiment can be seen in panel c of fig. 5.3, where evidence of travelling wave solutions can be seen in panel c.2. The invasion speed of the weak heterozygote invaded slower than that of both the intermediate and strong heterozygote.

Table 5.4: Average invasion speeds at different rates of mutation when the heterozygote expresses strong, intermediate, and weak traits of each allele.

Heterozygote traits	**Mutation parameter**	**Average speed**
Strong heterozygote $r_{AB} = r_{AA}$, $D_{AB} = D_{BB}$	$\mu = 0$	0.8501
	$\mu = 1\text{x}10^{-6}$	0.8501
	$\mu = 0.001$	0.8501
	$\mu = 0.01$	0.8506
	$\mu = 0.1$	0.8506
Intermediate heterozygote $r_{AB} = 2.55$, $D_{AB} = 0.25$	$\mu = 0$	0.5377
	$\mu = 1\text{x}10^{-6}$	0.5377
	$\mu = 0.001$	0.5380
	$\mu = 0.01$	0.5403
	$\mu = 0.1$	0.5576
Weak heterozygote $r_{AB} = r_{BB}$, $D_{AB} = D_{AA}$	$\mu = 0$	$c_1 = 0.0740$, $c_2 = 0.2197$
	$\mu = 1\text{x}10^{-6}$	$c_1 = 0.1914$, $c_2 = 0.2197$
	$\mu = 0.001$	0.2195
	$\mu = 0.01$	0.2209
	$\mu = 0.1$	0.3034

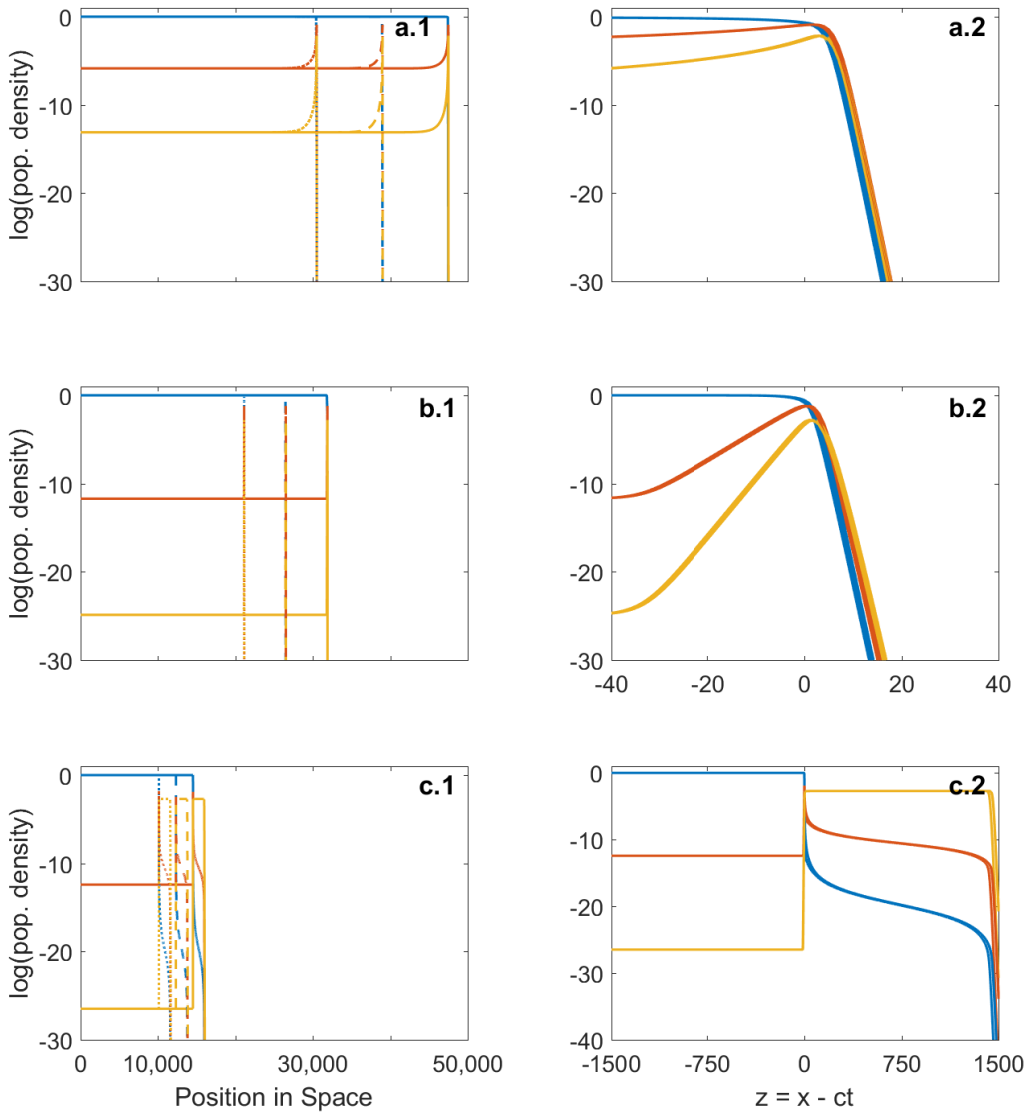

Figure 5.3: Invasion of vacant habitat by all genotypes. The natural log. of population density is given on the y-axis. Position in space is given on the x-axis. the reproductive genotype is shown in blue, the dispersive genotype is shown in orange, and the heterzygote is shown in yellow. Parameter values and initial conditions are given in table 5.3 and table 5.4 with $\mu = \text{x}10^{-6}$. Panels row a show the invasion when the heterozygote has the strongest traits of both homozygotes. Panels row b shows the invasion when the heterozygote has intermediate traits of both homozygotes. Panels row c shows the invasion when the heterozygote has the weakest traits of both homozygotes. Panels a.1, b.1, and c.1 depict invasion fronts at 3 equidistant time shots at $t = 30,000$ (solid line), $t = 40,000$ (dashed line), and $t = 50,000$ (dotted line). Panels a.2, b.2, and c.2 depict the natural log. of the population density against the wave variable $z = x - ct$ of each case at 10 time points, 5,000 time steps apart..

5.4 Discussion

We have developed a general model that is discrete in both time and space to investigate the invasion dynamics of a sexually reproducing diploid species. We then investigated a simplified case where there were no differences between males and females, but that individuals differed according to a dispersal-reproduction trade-off expressed through genotypes. We then examined the different possible cases when the heterozygote expressed traits from both the dispersive and reproductive homozygotes. Each case was then investigated through numerical experiments to establish the existence of travelling wave solutions, and whether anomalous invasion speeds were realised. We then drew comparisons to our previous chapters where mutation was critical to the existence of anomalous invasion speeds by investigating differing rates of mutation from one allele to another, i.e. a mutation at birth from A into B, and *vice versa*.

In all cases we found that travelling wave solutions were realised. We established that anomalous invasion speeds occurred when one allele was dominant causing the heterozygote to express the same traits as either homozygote – except in the case when reproduction was dominant and there was no mutation between alleles where we witnessed a stacked front. Upon examining the cases when the heterozygote expressed a mix of the homozygotes' traits, we found that anomalous invasion speeds were seen when the heterozygote had intermediate trait values for reproduction and dispersal, and only when mutation was high enough when it had a combination of the weakest traits of both homozygotes. We can conclude that when the heterozygote had a combination of the strongest traits of the homozygotes that anomalous speeds were not realised because, recalling from chapter 2, the shape of the dispersal-reproduction curve had negative curvature and thus the population would invade at the optimum speed. When the heterozygote expresses the strengths of both homozygotes is analogous to the aforementioned case.

We witnessed that, unlike in chapters 2 & 3, mutation played a less important role in the expression of anomalous speeds, except in two noteworthy cases.

We found that stacked fronts occurred when the dispersal abilities of both the heterozygote and homozygote were not large enough to keep up with the invasion led by the dispersive homozygote, this of course changed when mutation between alleles was large enough. We can conjecture that when the reproductive abilities of the disperser are weak that it can be invaded by the other genotypes with stronger reproductive fitness, which would explain the stacked fronts we have seen. For the population to invade at a single speed – and by extension, anomalous invasion speeds – the reproductive homozygotes and heterozygotes must keep pace with the invasion and are able to do so when there is enough exchanging of genotypes, i.e. individuals of each genotype born at the wave front. In the cases when we witnessed stacked fronts there was not enough exchanging of genotypes, until mutation was beyond a threshold that allowed them to keep pace with the dispersive homozygotes. Therefore, we can develop a general notion that unless there is a mechanism for maintaining less dispersive individuals at the wavefront then the population cannot invade at the same speed. Moreover, the whole population cannot benefit from anomalous invasion speeds afforded to them by the traits that give rise to such speeds.

We have shown through the course of this chapter that anomalous invasion speeds persist with sexual reproduction, and are not a phenomena that is restricted to clonal populations. Understanding the dynamics of invasion is important for managers within conservation and population control of invasive species. The implications of our findings are that invasions could be happening at a faster rate than current predictions have forecast, in this study we found a speed up of approximately a factor of 2 (see tables 5.3 & 5.4). Although explanations for some faster invasions have firm grounding in the evolution of dispersal at the wavefront, such as in cane toads (*Rhinella marina*) in Western Australia [136], but that is not to say that anomalous invasion speeds are not actively realised within such systems.

Chapter 6

Discussion

6.1 General Discussion

In a world that will be increasingly dominated by climate change, species will respond by adjusting their ranges to the changes, or perish if they cannot adapt to the changing conditions. The impacts of these changes will have wide ranging consequences for species globally, including our own. We will experience economical and health challenges related to expansion and invasion, as well as moral challenges through our responsibility for collective pressure on the natural world. Global biodiversity is predicted to decrease as a result of climate change and many species are expected to be negatively affected by these changes [175, 157, 9]. Understanding how quickly a species advances is of critical importance if we want to mitigate the impacts of invasive species, or to determine if a species can keep up with the pace of a shifting climate.

Biological invasions are developing at increasing rates because of human activity, by acting as vectors, or putting pressure on the environment resulting in removing other species that act as competitive barriers [156]. The effects on local ecosystems can be devastating, as in Western Australia from the cane toad invasion where the natural toxins produced by the toads are increasing mortality rates

among their non-natural predators [160]. The ability of non-native invaders to disrupt the ecosystem can have huge implications on human economic interests, such as increasing fire risk [37], or threatening food production [206], or introducing diseases [50]. Tracking the rate of invasion can then allow us to mitigate the damages caused by implementing strategies to combat the negative effects brought about by expanding or invading species.

Mathematical models have a long history of developing our understanding of biological processes, particularly those where experiments are seemingly impractical and costly. This is especially true for understanding range expansions and invasions when the scale is large, and uncertainty of empirical measurements are high. There is ample and growing evidence that dispersal evolves during the expansion process [137, 175, 181], the consequences of which are only beginning to be understood. The trend of all investigations into the evolution of dispersal is to find faster invasions than if dispersal does not evolve [136, 55]. Dispersal cannot evolve indefinitely, so there must exist a maximum energy allocation, that as it increases, some other physiological process will suffer in the form of a trade-off. Few studies have examined trade-offs, but those that do indicate that the population dynamics are altered in a non-negligible way [47, 48, 17].

Throughout this thesis we have shown that including dispersal trade-offs and dispersal polymorphism affects population dynamics during expansion by facilitating faster invasions through anomalous speeds than if any strain were in isolation. We shall now recall the key results of the thesis and draw the overarching conclusion that dispersal polymorphism and dispersal trade-off are properties of life history that have profound consequences for any predictions about invasion. Furthermore, we build on the theory regarding anomalous invasions speeds, providing further evidence that this phenomenon is not an artefact of the models they were studied in, but a robust effect of dispersal polymorphism and dispersal trade-offs.

6.1.1 Dispersal trade-offs

The evidence for dispersal trade-offs within the literature has been synthesised through narrative reviews [102, 12, 22, 155], but never before has it been quantified. The problem, as we have stated earlier, with narrative reviews is that they can miss some of the key findings, and that evidence may exist outside of the traditional areas of research. In chapter 4 we conducted a synthesis of the literature using a systematic review with broad search terms to capture as much of the studies as possible to use within a meta-analysis of dispersal trade-offs. We found that the evidence for dispersal trade-offs does exists.

We were able to achieve a quantified distribution of dispersal types, fitness types, species classes, sex of the studied organisms, climate classification, and environmental lineage within the literature. This is an excitingly useful database for any researcher that is interested in dispersal, as well as trade-offs. The meta-data we have found can now be utilised in other quantitative studies, or to guide empirical studies searching for dispersal trade-offs. Until now, the apparent absence of any trade-offs has left theoretical work unverified. We now have further evidence that dispersal trade-offs exist giving theoretical studies, and predictions relying on them, increased credibility.

We found significant trade-offs when we investigated the variation within the effect sizes. The factors responsible for these trade-offs were found to be fecundity, if the organism was female, if the organism was from a temperate climate, and if the organisms were insects. Additionally when we examined the type of dispersal, it was when a physical investment into dispersal apparatus was made that a trade-off became apparent. Some of our results are not surprising, such as trade-offs between dispersal and fecundity, within females, and within insects. Theoretical work predicts that investment in reproduction at the expense of dispersal should exist [96, 192], and empirical studies have shown this to be true too – with a fantastic meta-analysis by Guerra in crickets [66, 86]. Much of our data was from crickets, but that does not take away from the other Insecta that the study included. Females are also expected to invest heavily in reproduction, and thus

we should expect to see this too. However, when we controlled for these factors, the trade-off still existed indicating that the factors we would expect correlation in were not proxies for the others. Therefore, we can draw the conclusion that there were true biological effects.

What was surprising was to find that individuals from temperate climates experienced a dispersal trade-off. We can surmise that the relatively moderate climatic conditions do not selectively disadvantage dispersal trade-offs, whereas the relatively harsh conditions of say polar, dry, and continental climes with large variation in conditions, select against trade-offs. We refrain from speculating on tropical climates because there were only 3 studies within the largest data set.

An important result was that we found no difference between individuals that were reared for some generations within laboratory stocks, and those collected directly from their natural environment. A general concern when investigating matters of trade-offs is whether the conditions within the laboratory organisms are distorting results. Inadvertently, or deliberately, they are under different selection pressures from those in the wild. Our results show that conditions do not seem to affect trade-offs, which is an important finding. This implies that laboratory stock can continue to be used to investigate dispersal trade-offs which can save resources that would otherwise be used to conduct experiments from wild organisms.

6.1.2 Dispersal-reproduction trade-offs & anomalous invasion speeds

Our investigations in chapters 2,3, and 5 were theoretical studies of invasions of species that experienced a dispersal-reproduction trade-off. There have been several studies that examined theoretical populations that allowed either dispersal, or reproduction to vary, but few that vary together, and fewer still explain the biological motivation to do so [194, 193, 134, 61, 42]. We have expanded on the theory developed by Elliott & Cornell for dispersal polymorphism and its effects during invasions to a wider family of cases [47, 48]. We have seen that when the

degree of polymorphism is expanded to the general N-strain case the resultant invasion rate is determined by at most two constituent strains. Through our analysis and numerical experiments we found that the predicted linear speeds and those calculated from the simulations were in good agreement, allowing us to conclude that the speeds were the same – which also agrees with the rigorous mathematical proof conducted by Girardin [61]. Furthermore, that the predicted speed of a 5-strain species invaded at the same speed as a 2-strain species provided the same vanguard phenotypes were present. When mutation dynamics were altered to a more realistic regime of small progress to either neighbouring discrete strain the result was unchanged.

A natural extension to the theory was to consider variation in how strains competed among themselves, and against another resident species. Studies which examine differences in competition are uncommon, often density dependence is included, but that interactions are neutral, and fewer examine interspecific competition. Those few that did investigate the effects of competition between species generally found that invasions were either slower, or halted altogether [194, 108, 129]. When we investigated intraspecific competition between strains invading vacant habitat we seen that anomalous invasion speeds were still possible, even when the vanguard strains were out-competed within the range core. The vanguard strains were present at the range front, driving the invasion forward, but when the range core catches up they are reduced to low densities. Anomalous invasion speeds were also found when a polymorphic species invaded a resident, albeit the effects of competition slowed the rate of advance, as has been seen in previous studies. It was here that we found the surprising results that the polymorphic species could still invade anomalously, even when the effects of competition caused one of the vanguard strains to experience negative growth. The conditions for which either monomorphic, or the polymorphic species could invade each other was determined by the effects of the competitor on the focal species, and that it's effect on the competitor did not affect if it could invade.

In chapter 5 we provide evidence that anomalous invasion speeds are robust

to a change in reproductive system from clonal to sexual. Using a discrete time and space difference model we conducted numerical experiments on a diploid population that experienced a dispersal reproduction trade-off through the expression of alleles. Our homozygotes either had strong dispersal, and weak reproduction, and *vice versa*. In doing so we could test how the heterozygote affected the invasion when it's expression of traits varied. If the heterozygote experienced dominance in either allele we found that anomalous invasion speeds were realised. If the heterozygote expressed the strengths of both alleles the invasion was fastest, and slowest if it expressed the weakness of both alleles, and an intermediate speed with intermediate expression.

Previous theoretical studies have shown that the shape of the trade-off curve can have serious implications for population dynamics [192, 96, 151]. However, few have applied this to the theory of range expansion and invasions [17, 47, 48]. We have shown that the shape of the trade-off curve is paramount for determining whether anomalous invasion speeds occur. If the curve has positive curvature, and the dispersal and reproductive abilities are sufficiently different, then anomalous invasion speeds will occur. If the trade-off curve has negative curvature then we see that a single strain will determine the invasion speed.

6.2 Implications & future directions

Through the course of this thesis we have found results that have wide reaching implications for research into range expansions and invasions. Theoretically we have shown that faster, anomalous invasion speeds are persistent in a variety of scenarios, which suggests that this phenomenon is not an artefact of a particular system, but a general behaviour that arises from the dynamics.

To determine if an anomalous speed is present requires knowing the dispersal-reproduction trade-off. Acquiring an accurate estimate of the trade-off curve is currently very difficult due to the complexities of measuring both the dispersal and reproduction traits effectively. Further developments in techniques that could

measure the traits required for our predictions is a must, since without them, we cannot validate our theories. A starting point would be to identify candidate species that display the kind of dispersal-reproduction trade-off to perform empirical experiments with. Our work from chapters 2 and 3 are within haploid species which reproduce clonally, so empirical tests on similar organisms from microbiology could provide fruitful. However, within microbes a dispersal reproduction trade-off would be difficult to find, or to create selection lines that display one. The fitness properties of microbes, and any trade-offs, are unlikely to sacrifice reproduction before any other trait (Goodman, Gupta, Williams, personal communication). We have however shown anomalous speeds persist when reproduction is sexual, where dispersal-reproduction trade-offs and candidate species do exist, like the cone-headed bush cricket which is expanding its range [175].

Our chapter on anomalous invasion speeds with sexual reproduction has shown using simulations the possibility of their existence, but not using mathematical analysis. Developing analytical methods for this system, as we have done with chapters 2 and 3, would provide greater insights to the dynamics. The possibility of anomalous invasion speeds for a polymorphic species invading vacant habitat has been proven mathematically by Girardin [61], but no current proof exists for when a competitor is present. To develop a proof for the case when an established competitor is present would give our work further weight and more justification in finding empirical candidate systems.

Within our models we have assumed that dispersal is a simple process of small local movements throughout the life time of an individual, but dispersal in nature is more complex. Further work to establish the theory for types of dispersal other than diffusion would be beneficial, particularly when real data is available. A new project called *ICARUS* has been establish by the Max Planck Institute for Ornithology which has sensors mounted on the International Space Station which can globally track the dispersal, movement, and migration patterns of animals [2]. The data that will become available can aid future dispersal research by parameterising models, testing dispersal theory, and deepening our understanding

of how organisms move more broadly. With more accurate information about dispersal we can then predict the rate of spread more accurately.

We have found that when there is general degree of polymorphism that the invasion speed is determined by two strains. Furthermore, when competition is accounted for we found that anomalous invasion speeds persisted when the vanguard strains were out-competed in the range core; and when one of the strains had negative growth from competitive pressure from another species. Both of these scenarios imply that predicting the invasion from empirical data would be problematic, since the vanguard strains would only be present in low densities. The implications for invasion management are even larger. Any attempts to intervene to slow the invasion effectively will be difficult to instigate correctly. Consider a highly polymorphic population invading anomalously. Population managers may not be aware that there is an anomalous invasion because a single strain is out-competing the rest in the core, believing the population to be monomorphic, but the vanguard strains are driving the invasion at the front. Another possibility is that a manager has predicted a slower invasion because a competitor is present. They have seen the decline of one of the vanguard strains as a result of the competitor and incorrectly interpret that an anomalous invasion is no longer occurring, when in fact, the invasion is still proceeding at the anomalous rate.

Clearly range expansions and biological invasions are complex eco-evolutionary processes that deserve scientific attention. As the increasing rate of climate change increases uncertainty for all organisms on the planet, understanding how species react to these conditions will better prepare humanity to mitigate the negative effects that can occur as a result. Loss of biodiversity is serious threat to the stability of global ecosystems, and through them, the stability of global economic and health systems.

Appendices

Appendix A

Geometric interpretation of anomalous speeds

A straight line passing through points representing strains i and j in (D, r) space has equation

$$r = r_i + (D - D_i)\frac{r_j - r_i}{D_j - D_i}, \tag{A.1}$$

so that it meets the axes at the points $(D, r) = \left(\frac{r_i D_j - r_j D_i}{r_j - r_i}, 0\right)$ and $(D, r) = \left(0, \frac{r_i D_j - r_j D_i}{D_j - D_i}\right)$. The midpoint of the line segment joining the two axes is then

$$(D, r) = (D_{vij}, r_{vij}) = \left(\frac{r_i D_j - r_j D_i}{2(r_j - r_i)}, \frac{r_i D_j - r_j D_i}{2(D_j - D_i)}\right), \tag{A.2}$$

and the monomorphic speed for this "virtual" strain is then $c = 2\sqrt{r_{vij} D_{vij}}$, which gives the same expression as c_{dij} from MS eq. (2). Thus, the dimorphic invasion speed for strains i and j is the same as the monomorphic speed for this virtual strain.

We assume (without loss of generality) that $r_i < r_j$, which implies that $D_j < D_i$ in order for the dimorphic speed in eqn. (3) to be real. The condition

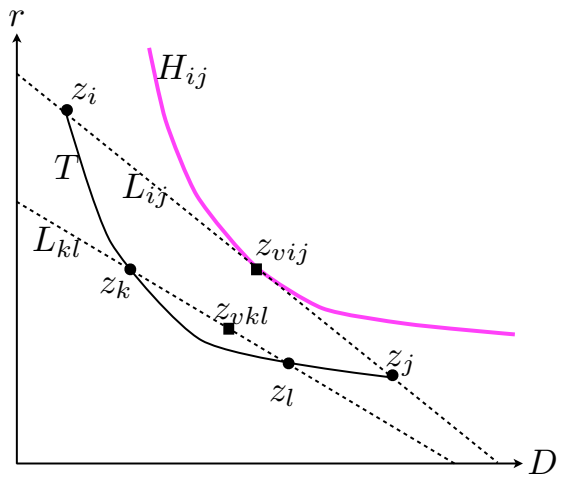

Figure A.1: The invasion speed when $D-r$ trade-off curve (represented by T) has positive curvature is the dimorphic speed for the virtual strain v_{ij} (represented by a point z_{vij}) for the two most extreme strains i and j (represented by points z_i and z_j), provided z_{vij} lies between z_i and z_j. H_{ij} is a hyperbola $r = \frac{D_{vij}r_{vij}}{D}$ that represents the locus of strains that would have the same monomorphic speed as v_{ij}; all strains represented by points below H_{ij} have $Dr < D_{vij}r_{vij}$, and therefore slower monomorphic speeds than the virtual strainv_{ij}. L_{ij} (resp L_{kl}) is a straight line segment passing through $z_i$and z_j (resp. z_k and z_l) and terminating on the r and D axes. As shown above, z_{vij} (resp. z_{vkl}) lies at the midpoint of L_{ij} (resp. L_{kl}). Since H_{ij} is tangential to L_{ij} at z_{vij}, the monomorphic speed for the virtual strain z_{vij} is faster than the monomorphic speed for any other strain on T. Moreover, since the virtual strain for any pair of strains k and l must lie between z_k and z_l for the corresponding dimorphic speed to be a valid invasion speed, so z_{vkl} must lie below L_{ij} and therefore below T, no dimorphic speed c_{dkl} can exceed c_{dij}.

for the virtual strain to lie between the real strains is then

$$r_i \leq r_{vij} \leq r_j,$$

and it is straightforward to show that this leads to exactly the same conditions as eqns. (4) in the MS. Moreover, it is straightforward to show that the virtual strain has the highest value of rD, and therefore the highest monomorphic speed $2\sqrt{rD}$, of any strain on the line segment defined by eqn. (A.1). This implies that the hyperbola

$$r = \frac{D_{vij}r_{vij}}{D}.$$

along which the monomorphic speed equals c_{dij}, is tangential to the line segment at $(D, r) = (D_{vij}, r_{vij})$

These geometric constructions allow us to find the invasion speed from the geometric properties of the trade-off curve alone. As shown in fig. A.1, if the curve segment representing the trade-off has positive curvature then the invasion speed is the anomalous dimorphic speed for the two most extreme strains, provided the corresponding virtual strain is bracketed by these extreme strains (i.e. that the condition eqn. (4) in the MS is satisfied so that this is a valid invasion speed). A similar construction shows that, if the trade-off curve has negative curvature, then no valid dimorphic speed is faster than the fastest monomorphic speed on the trade-off curve.

If the trade-off curve has segments with both positive and negative curvature (e.g. Figure 4 f) then the pair of strains that give the fastest dimorphic speed need not be at the extrema of the trade-off curve $r = r(D)$ provided there is a pair of strains with diffusion constants D_1 and D_2 where $\frac{\partial c_{d12}}{\partial D_1} = 0 = \frac{\partial c_{d12}}{\partial D_2}$. Taking derivatives of eqn. (3) gives

$$\frac{\partial c_{d12}}{\partial D_1} = \frac{\{D_1 r(D_2) + D_2 r(D_1) - 2D_2 r(D_2)\}\left[(D_1 - D_2)r'(D_1) + r_2 - r_1\right]}{2\left((D(r_1) - D(r_2))(D_2 - D_1)\right)^{3/2}} = 0,$$
$$\frac{\partial c_{d12}}{\partial D_2} = \frac{\{D_1 r(D_2) + D_2 r(D_1) - 2D_1 r(D_1)\}\left[(D_1 - D_2)r'(D_2) + r_2 - r_1\right]}{2\left((D(r_1) - D(r_2))(D_2 - D_1)\right)^{3/2}} = 0.$$

If either of the terms in braces is zero, then as seen in eqn. (A.2), the virtual strain is identical to either strain 1 or strain 2, so the dimorphic speed is the same as the monomorphic strains. If the dimorphic speed is truly anomalous, then the terms in braces must be non-zero, so we must have

$$r'(D_1) = r'(D_2) = \frac{r(D_1) - r(D_2)}{D_1 - D_2}.$$

The term on the right hand side is equal to the gradient of the straight line joining the two strains, so this line must therefore be tangential to the trade-off curve at both points (see Fig. 4 f).

To show that the vanguard morphs need not have the highest values of rD on the trade-off curve, we provide an illustrative example in figure A.2

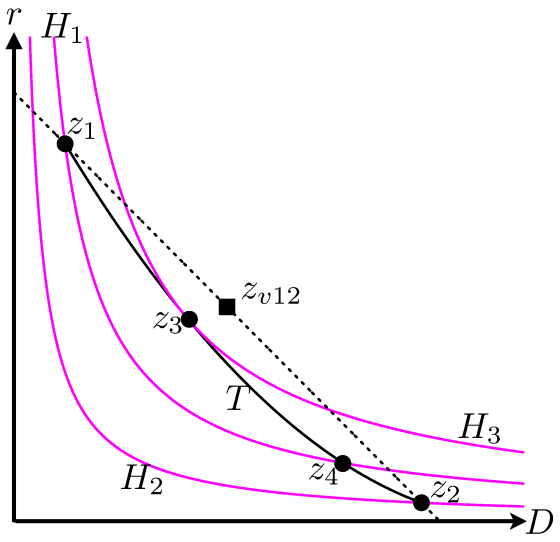

Figure A.2: Illustration that vanguard strains do not necessarily have the highest value of rD on the trade-off curve. Points z_1 and z_2 represent the strains at the endpoints of a convex trade-off curve (solid black line T) in (D, r) space, which as shown in fig. A.1 are the vanguard strains for a species with this trade-off curve. The dotted line is a straight line passing thorough z_1 and z_2, and z_{v12} represents the corresponding virtual strain constructed as in fig. A.1. Curves H_1 and H_2 are hyperbolae representing the loci of points where rD (and therefore the monomorphic speed) takes the same values as at z_1 and z_2 respectively. The segment of T between z_1 and z_4 (where T crosses H_1) lies above both H_1 and H_2, so all strains along this segment have a higher value of rD than either z_1 or z_2. In particular, H_3 is a hyperbola tangential to the trade-off curve at point z_3, so z_3 represents the strain with the highest value of rD on the trade-off. The virtual strain z_{v12} lies above H_3, in corroboration of the fact that the monomorphic speed for the virtual strain is higher than that for z_3, which has the highest monomorphic speed on the trade-off.

Appendix B

MATLAB code used for simulations

The script below is the general form of the code used to simulate the equations in chapters 2,3, and 5.

B.1 Chapters 2 & 3

B.1.1 CN matrix matrix1.m

```
%This is the function used to generate the matrix used in the first part
%of the Crank-Nicholson scheme. The matrix is a tridiagonal system with
%zero flux boundary conditions included.

%The main loop for creating the tridiagonal matrix. Each r(i) is different
%and must be accounted for - the loop takes this into consideration. Cell
%arrays are used as a simple way of producing block matrices. The B.C.s
%are also included here, in this circumstance they are zero flux.
function X=CNmatrix1(a,theta)
b=cell(1,length(a{3}));
for pp=1:length(a{3})
[b{pp}]=gallery('tridiag',length(a{1}),(1-theta)*a{6}(pp),...
1-2*(1-theta)*a{6}(pp),(1-theta)*a{6}(pp));
%Here we insert the boundary conditions which will carry over for each
```

```
%cell.
cellarray=[b{pp}];
cellarray(1,1)=1-2*(1-theta)*a{6}(pp);
cellarray(1,2)=2*(1-theta)*a{6}(pp);
cellarray(length(a{1}),length(a{1})-1)=2*(1-theta)*a{6}(pp);
cellarray(length(a{1}),length(a{1}))=1-2*(1-theta)*a{6}(pp);
[b{pp}]=cellarray;
end
%This step concatenates the cells into a block diagonal matrix.
X=(blkdiag(b{:}));
end
```

B.1.2 CN matrix2.m

```
%This is the function used to generate the matrix used in the second part
%of the Crank-Nicholson scheme. The matrix is a tridiagonal system with
%zero flux boundary conditions included.

%The main loop for creating the tridiagonal matrix. Each r(i) is different
%and must be accounted for - the loop takes this into consideration. Cell
%arrays are used as a simple way of producing block matrices. The B.C.s
%are also included here, in this circumstance they are zero flux.
function X=CNmatrix2(a,theta)
b=cell(1,length(a{3}));
for pp=1:length(a{3})
[b{pp}]=gallery('tridiag',length(a{1}),-theta*a{6}(pp),...
1+2*theta*a{6}(pp),-theta*a{6}(pp));
%Here we insert the boundary conditions which will carry over for each
%cell.
cellarray=[b{pp}];
cellarray(1,1)=1+2*theta*a{6}(pp);
cellarray(1,2)=-2*theta*a{6}(pp);
cellarray(length(b{pp}),length(b{pp})-1)=-2*theta*a{6}(pp);
cellarray(length(b{pp}),length(b{pp}))=1+2*theta*a{6}(pp);
[b{pp}]=cellarray;
end
%This step concatenates the cells into a block diagonal matrix.
X=(blkdiag(b{:}));
end
```

B.1.3 kinetics.m

```
%This is a function for calculating the kinetics to be inserted into the
%main calculation loop. I take each individual term within the kinetics and
```

```
%calculate these seperately, and then concatenate them at the end. This is
%the function for universal mutation.

function B=kinetics(a,dt)

%Initialise the arrays to be used within the loop.
b=cell(3,1);
c=cell(1,1);
d=cell(length(a{2}));
%This cell is for storing the mutation matrix with 0 on the diagonals for
%calculating the final terms within the kinetics easier.
c{2}=a{5}-diag(diag(a{5}));
%This loop calculates the kinetics for each individual species.
for ii=1:length(a{2})
%This term is the competition coefficient part.
b{1}=(-1)*sum(bsxfun(@times,a{4}(ii,:),a{1}),2);
%This term is the positive incoming mutations.
b{2}=sum(bsxfun(@times,c{2}(:,ii)',a{1}),2);
%This term is the negative leaving mutations.
b{3}=(-1)*sum(c{2}(ii,:))*a{1}(:,ii);
%This term is the full kinetic equation using the above.
d{ii}=(a{2}(ii).*a{1}(:,ii).*(1+b{1})+b{2}+b{3})*dt;
end
%Concatenating the matrix into the form used for solving numerically.
B=cat(1,d{:});

end
```

B.1.4 nnkinetics.m

```
%This is a function for calculating the kinetics to be inserted into the
%main calculation loop. I take each individual term within the kinetics and
%calculate these seperately, and concatenate them at the end. This is the
%function for nearest-neighbour mutation.

function B=kinetics(a,dt)

%Initialise the arrays to be used within the loop.
b=cell(3,1);
d=cell(length(a{2}));
%This cell is for storing the mutation matrix with 0 on the diagonals for
%calculating the final terms within the kinetics easier.
mu=(a{5}(1,2))/2;

%Creates empty vectors for use in the kinetic terms.
```

```
for i=1:3
b{i}=zeros(length(a{2}),1);
end

%This loop calculates the kinetics for each individual species.
for ii=1:length(a{2})
% This term is the competition coefficient part.
b{1}=-1*sum(bsxfun(@times,a{4}(ii,:),a{1}),2);
%These terms are the incoming and leaving mutations for each species.
if ii==1                          %case for 1 due to indexing
b{2}=a{1}(:,2)*mu;
b{3}=-1*a{1}(:,1)*mu;
elseif ii==length(a{2})           %case for N due to indexing
b{2}=a{1}(:,end-1)*mu;
b{3}=-1*a{1}(:,end)*mu;
else
b{2}=(a{1}(:,ii-1)+a{1}(:,ii+1))*mu;
b{3}=-1*a{1}(:,ii)*2*mu;
end
%This term is the full kinetic equation using the above.
d{ii}=(a{2}(ii).*a{1}(:,ii).*(1+b{1})+b{2}+b{3})*dt;
end
%Concatenating the matrix into the form used for solving numerically.
B=cat(1,d{:});

end
```

B.1.5 wavefront.m

```
%%xwavefront.m is a piece of code designed for finding the wavefront of
%%travelling wave solution. Often there are more than one value that are
%%numerically similar to the wavefront. This code identifies the
%%wavefrontand finds the midpoint of the wave to take a measurement of the
%%speed. This assumes that the population peaks at 1.
function X=xwavefront(a)
%finds the maximum value in the solution - this is just prior to the
%wavefront. The output is the value found maxNum, idx is the index of the
%value.
[~, idx]=max((a{1}(:)));
%u is a vector from the max solution value to the end of the original
%solution.
u=a{1}(idx:end);
%This finds the midpoint of the wave by finding the value closest to 0.5.
%val is the value, idx is the matrix index.
[~,idx]=min(abs(0.5-u));
```

```
%This finds the value in the original solution.
[row, col]=find(a{1}(:)==u(idx));
%This changes the linear indexing of the value to the more familiar
%(row,column) indexing which is what we need.
[row, col]=ind2sub(size(a{1}),row);
%The output is then the row where this value is which corresponds to
%spatial value pinpointing the wavefront.
X=max(row);
end
```

B.1.6 nmorphsol3.m

```
%Function for solving the solution of the reaction diffusion equations. The
%function uses the Crank-Nicholson scheme. The kinetic terms should be
%adjusted for the users need. The input is a cellular array, the input
%for the a{1} must be an nxn matrix.
function X=nmorphsol3(a,dt,theta,dx,Lmax)

%allocates the Crank-Nicolson matrices used for the calculation.
a{7}=CNmatrix1(a,theta);
a{8}=CNmatrix2(a,theta);
%Numerical solver which obtains the solution.
[L,U,P] = ilu(a{8},struct('type','ilutp','droptol',1e-6));
%Generate cells and vectors for use in the function.
b=cell(3,1);

k=1;        %Creates a count for saving the solution at time intervals.
a{12}=zeros(101,2);

%hvec is used later for a solution shift in space.
hvec=length(a{10}(floor(length(a{10})/2):end));

%% Main solution loop

for ii=1:length(a{11})
%Compresses the kinetics calculations into an array
a{9}=kinetics(a,dt);
%First step of the Crank-Nicolson scheme
x=a{7}*a{1}(:)+a{9};
%Second step of the Crank-Nicolson scheme. Using an LU decomposition
%for faster calculation.
q=U\(L\(P*x));

%This part of the code has been included in an attempt to remove the
%problem of small numbers within the solution propagting changes in
```

```
%decay rate through the solution due to round-off error.
q(q < realmin)=0;
%Reshapes the solution into a matrix corresponding to each strain.
a{1}=reshape(q,length(a{1}),length(a{2}));

%Finds the midpoint of the wave.
xrow=xwavefront(a);
%IF the midpoint is 9/10*L then the solution is shifted by a half. This
%keeps the boundary the same size, but shifts space for a continuing
%solution. This is a step to reduce computation time.
%     if xrow(end) >= 9*(length(a{10})/10)
%        %This part shifts the solution.
%        nzeros=zeros(size(a{1}));
%        nzeros(1:length(hvec:end),:)=a{1}(hvec:end,:);
%        a{1}=nzeros;
%        %This part shifts the space.
%        a{10}=a{10}(hvec):dx:(a{10}(hvec)+hvec*dx);
%     end

%If loop for saving equally spaced time points throughout the
%simulation.
if rem(ii,floor(length(a{11})/100))==0
%Here we store the wavefront at it's time and position within the
%simulations in a{12}.
xrow=xwavefront(a);
a{12}(k,1)=a{10}(xrow);
a{12}(k,2)=a{11}(ii);
%Seperate array for saving the essential information required for
%assessing speeds, i.e. the numerical solution, b{1}, at the time
%and position, b{2}.
b{1}=a{1};
b{2}=a{12};
b{3}=a{10};
%Solution and plot from the solution is saved.
save(['//home/vakeenan/paperfigs2/5morph/sim',num2str(k),'at_x_',num2str(a{10}(xrow)),...
'_t_',num2str(a{11}(ii)),'.mat'], 'b')

k=k+1;
end

%IF the midpoint is 9/10*L then the solution is shifted by a half. This
%keeps the boundary the same size, but shifts space for a continuing
%solution. This is a step to reduce computation time.

%     if xrow(end) >= 9*(length(a{10})/10)
%        %This part shifts the solution.
%        nzeros=zeros(size(a{1}));
```

```
%        nzeros(1:length(hvec:end),:)=a{1}(hvec:end,:);
%        a{1}=nzeros;
%        %This part shifts the space.
%        a{10}=a{10}(hvec):dx:(a{10}(hvec)+Lmax);
%     end

% There are two shifts, both commented out. This is because the user may
% want the shift done before or after saving and is up to them.
end

X=a;


end
```

B.1.7 ElliottCornellequations.m

```
% MATLAB code for solving Elliott & Cornell equations in 1-D for any number of morphs
% using a Crank-Nicholson scheme. Composed by Vincent Keenan. All of the
% calculated values are stored in an array "a". The numerical allocation of
% the array is 1=n, 2=R, 3=D, 4=C, 5=MU, 6=r, 7=CNmatrix1, 8=CNmatrix2,
% 9=kinetics, 10=x, 11=t, 12=x and t values for approximating the speed.

%% Initialise mesh
%L is the length of space which acts as the environment. T is the time
%over which the simulation models.
tic                   %Begins timer to determine run time.
L=28000;
Lmax=L;
T=7200;
a=cell(13,1);
%temporal and spatial step sizes.
dx=0.05;
dt=0.05;
%creating vectors for time and space to be used within the calculations.
%Other parameters are included for ease later in the calculation.
a{10}=(0:dx:L);      %Intialise space discretisation.
a{11}=(0:dt:T);      %Intitialise time discretisation.

%% Initialise parameters
%parameters of the equation which can be selected by the user. Ri are
%reproductive parameters. Di are dispersal parameters. MUi are mutation
%parameters. Theta is set at 1/2 for use within the Crank-Nicholson scheme
%for solving the PDE.
theta=1/2;
```

```
D=[8.5, 8, 4.5, 3, 0.5];          %Vector containing diffusion constant.
R=[0.05, 0.3, 0.45, 0.8, 0.85];   %Vector containing growth rates.
a{2}=R;
a{3}=D;
MU=0.005*ones(length(a{2}));      %Matrix containing mutation rates.
C=ones(length(a{2}));             %Matrix containing competition coeff.
a{4}=C;
a{5}=MU;

%Vector that contains r(i) values that are important for stability of
%finite difference methods - analagous to the CFD condition. For the
%Crank-Nicholson method this is not an issue as it is always stable for all
%values of r(i). Not to be confused with per-capita growth rate. See
%literature for numerical methods if unsure.
a{6}=zeros(length(a{2}),1);
for ii=1:length(a{2})
a{6}(ii)=(a{3}(ii)*dt)/(dx^2);
end



%% Initialise Initial conditions
%Creating the initial vectors for each morph and concatanating them for
%ease within the calculation.
a{1}=zeros(length(a{10}),length(a{2})); %Solution vector.I'C's overwritten.

%Loop that fills the matrix with the initial conditions for the specified
%spatial location and pop. density values.

for jj=1:(L/dx/100)          %spatial entries to be filled.
a{1}(jj,:)=1/length(a{2});   %pop density values
end


%% Main solution loop for solving the equation.

a=nmorphsol3(a,dt,theta,dx,Lmax);

%% Data Extraction

%For calculating the speed of the travelling wave using "v=d/t".
space=diff(a{12}(1:end-1,1));
time=diff(a{12}(1:end-1,2));
speed=space./time;

%Here we store all the appropriate information in b{} required for
```

```
%reporting in the manuscript. Alternatively for all values the user could
%choose to save a{} instead.
a{13}=mean(speed);
b=cell(5,1);
b{1}=a{1};
b{2}=a{12};
b{3}=a{13};
toc;
soltime=toc;
b{4}=soltime;
b{5}=a{10};
save(['Enter your directory here/final_at_x _',num2str(a{12}(end,1)),...
'_t_',num2str(a{11}(end)),'.mat'], 'b')
```

B.2 Chapter 5

B.2.1 diploidinvmod.m

```
%% Diploid polymorphic species invasion model
% Here is the discrete time and space model used to simulate an invasion of
% a diploid polymorphic species invading empty habitat. The species has
% genotypes AA, AB, and BB. The genotype AA is the high D, low r phenotype;
% the BB is the low D, high r phenotype; and the heterozygote is somewhere
% between both. The fitness, encapsulated within the per capita growth
% rate. The diploid_temp_model.m must be run first to determine the
% non-trivial equilibrium value to normalise the densities.

%% Preamble
% Here we have some arrays which need to be preallocated before they are
% used for storage of outputs.

%Here we store the total number of timepoints we want to save over.
total_tp = 100;

%create a vector to store the position of the wavefront at time timepoint.
wavefront = zeros(total_tp, 6);

%create a vector to store the timepoint which to save the wavefront
%position.
timepoint = zeros(total_tp,1);

%Createa  counter to store outputs at specified timepoints.
k = 1;
```

```
%% Initial conditions

%Initialise spatial and temporal dimensions.
L = 1:50000; T = 1:50000;

%Generate vectors for each gendered genotype for each position in space.
f_AA = zeros(length(L), 1); m_AA = zeros(length(L), 1);
f_AB = zeros(length(L), 1); m_AB = zeros(length(L), 1);
f_BB = zeros(length(L), 1); m_BB = zeros(length(L), 1);

%Include an intial condition.
f_AA(1:(length(L)/10),1) = 1; m_AA(1:(length(L)/10),1) = 1;
f_AB(1:(length(L)/10),1) = 1; m_AB(1:(length(L)/10),1) = 1;
f_BB(1:(length(L)/10),1) = 1; m_BB(1:(length(L)/10),1) = 1;

%Concatenate gendered gentotypes into a single matrix for each position in
%space.
N_tot = [f_AA, m_AA, f_AB, m_AB, f_BB, m_BB]*1/6;

%per-capita growth parameters. Implicit fitness attached to female
%reproductive success.
r_AA = 5; r_BB = 0.1; r_AB = 0.1;



%% Main time loop
%This is the main soluation loop.
for i = 1:length(T)
%% Mating and growth dynamics including mutation
%growth dynamics are given here.

%Reproductive combinations of different genotypes by gender, i.e. A_f is
%the female contribution to allele A, etc. Females have column vectors.
A_f1 = [r_AA * N_tot(:,1), (1/2) * r_AB * N_tot(:,3)];
B_f1 = [r_BB * N_tot(:,5), (1/2) * r_AB * N_tot(:,3)];

%Probability of encountering a male with genotype A or B. Males have row
%vectors.

Z = zeros(length(N_tot),1);

for j = 1:length(N_tot)
if N_tot(j,2) + N_tot(j,4) + N_tot(j,6) == 0
Z(j) =0;
else
Z (j) = 1 ./(N_tot(j,2) + N_tot(j,4) + N_tot(j,6));
end
end
```

```
A_m1 = bsxfun (@times, Z, [N_tot(:,2), (1/2) * N_tot(:,4)]);
B_m1 = bsxfun (@times, Z, [N_tot(:,6), (1/2) * N_tot(:,4)]);


%Mutation parameter that controls the frequency of mutations in any allele.
mut.par = 1e-6;

%New populations of genotypes after mutation. 1-mut.par of the pop remain
%as original genotype, and mut.par of the pop change genotype. We must take
%the transpose when adding the mutated alleles to match dimensions.
A_f = A_f1 * (1 - mut.par) + B_f1 * mut.par;
B_f = B_f1 * (1 - mut.par) + A_f1 * mut.par;
A_m = A_m1 * (1 - mut.par) + B_m1 * mut.par;
B_m = B_m1 * (1 - mut.par) + A_m1 * mut.par;

%Birth rates of each individual genotype. Here we take the matrix product
%and sum over all entries to give us the growth rate.
R_AA = sum(bsxfun(@times, A_f(:,1), A_m) + bsxfun(@times, A_f(:,2),A_m),2);
R_AB = sum(bsxfun(@times, A_f(:,1), B_m) + bsxfun(@times, A_f(:,2),B_m)...
+ bsxfun(@times, B_f(:,1), A_m) + bsxfun(@times, B_f(:,2),A_m),2);
R_BB = sum(bsxfun(@times, B_f(:,1), B_m) + bsxfun(@times, B_f(:,2),B_m),2);

%Here we add the births to the population. The proportion of females
%birthed is alpha. This could be perturbed stochastically if desried.
alpha = 1/2;
N_tot_new = zeros(size(N_tot));
N_tot_new(:,1) = N_tot(:,1) + (alpha * R_AA);
N_tot_new(:,2) = N_tot(:,2) + ((1 - alpha) * R_AA);
N_tot_new(:,3) = N_tot(:,3) + (alpha * R_AB);
N_tot_new(:,4) = N_tot(:,4) + ((1 - alpha) * R_AB);
N_tot_new(:,5) = N_tot(:,5) + (alpha * R_BB);
N_tot_new(:,6) = N_tot(:,6) + ((1 - alpha) * R_BB);

%Here we calculate the total population for the density dependent stage.
dn_N_tot = (0.8333) * sum(N_tot_new,2);

%% Density dependent mortality
%This accounts for the term within the denominator of the the Leslie-Gower
%competition terms, provided they are neutral.

%Density dependent growth for AA male and female genotypes.
N_tot(:,1) =  N_tot_new(:,1) ./ (1 + dn_N_tot);
N_tot(:,2) =  N_tot_new(:,2) ./ (1 + dn_N_tot);

%Density dependent growth for AB male and female genotypes.
N_tot(:,3) = N_tot_new(:,3) ./ (1 +  dn_N_tot);
```

```
N_tot(:,4) = N_tot_new(:,4) ./ (1 +  dn_N_tot);
%Density dependent growth for AA male and female genotypes.
N_tot(:,5) = N_tot_new(:,5) ./ (1 +  dn_N_tot);
N_tot(:,6) = N_tot_new(:,6) ./ (1 +  dn_N_tot);


%% Dispersal
%Here we define dispersal.

%Here is the proportion which disperse within the population.
D = [0.00055; 0.00055; 0.5; 0.5; 0.5; 0.5];

disp_N = zeros(length(N_tot),length(D));

%total proportion dispersing.
for ii = 1: length(D)
disp_N(:,ii) = D(ii) * N_tot(:,ii);
end

% This is the dispersing individuals. A proportion reamin, 1 - D_i, while a
% proportion disperse to neighbouring cells, D_i.
for ii=1:length(N_tot)

if ii==1                          %case for 1 and 2 due to indexing
N_tot(1,:) = N_tot(1,:) + (1/2) * disp_N(2,:) ...
- (1/2) * disp_N(1,:);
elseif ii==length(disp_N)           %case for N due to indexing
N_tot(end,:) = N_tot(end,:) + (1/2) * disp_N(end-1,:) ...
- (1/2) * disp_N(end,:) ;
else
N_tot(ii,:) = N_tot(ii,:) + (1/2) * (disp_N(ii-1,:) ...
+ disp_N(ii+1,:)) - disp_N(ii,:);
end

end

%We set this condition to remove floating point errors that creep into the
%calculation during the dispersal phase.
N_tot(N_tot < realmin)=0;


%Here I set out a condition to save at specified timepoints throughout the
%simulation.
if rem(i, floor (length (T) / total_tp)) == 0
xrow = zeros(size(N_tot,2), 1);
for ii = 1:size(N_tot,2)
```

```
[~, idx] = max (N_tot (:,ii) ) ;

%u is a vector from the max solution value to the end of the
%original solution.
u = N_tot (idx:end, ii);

%This finds the midpoint of the wave by finding the value closest
%to 0.5. val is the value, idx is the matrix index.
[~,idx] = min (abs ((max(u)/2) - u) );

%This finds the value in the original solution and provides a
%location in space to each phenotype.
[row, ~] = find (N_tot (:, ii) == u (idx) );
xrow(ii) = max (row);
end

%Here we store the wavefront at it's time and position within the
%simulations in a{12}.

for  J=1:size(N_tot,2)
wavefront(k,J) = L(xrow(J));
end

timepoint(k) = T(i);

%Seperate array for saving the essential information required for
%assessing speeds, i.e. the numerical solution, b{1}, at the time and
%position, b{2}.
b{1} = N_tot;
b{2} = wavefront;
b{3} = timepoint;
b{4} = L;
b{5} = L - 0.45 * T(i);

%Solution and plot from the solution is saved.
save(['//home/vakeenan/diploid_inv/not_extreme/mue6/heteroD0/sim',num2str(k),...
'at_t_',num2str(T(i)),'.mat'], 'b')

k=k+1;
end
end
% plot(L,N_tot)

space=diff(wavefront);
time=diff(timepoint);
speed=bsxfun(@rdivide, space, time);
```

```
b{1} = N_tot;
b{2} = wavefront;
b{3} = timepoint;
b{4} = L;
b{5} = L - 0.45 * T(i);
b{6} = speed;

save(['//home/vakeenan/diploid_inv/not_extreme/mue6/heteroD0/final_at_x _',...
num2str(max(wavefront(end,:))),'_t_',num2str(T(end)),'.mat'], 'b')
```

Appendix C

Diagnostics example

There were over 120 models run, each of which would require appropriate diagrams. Here we show the example for the model

```
ES.model.ran_artID <- MCMCglmm(SMD_g ~ 1,  random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE).
```

All other diagnostics were considered ok and looked similar to these.

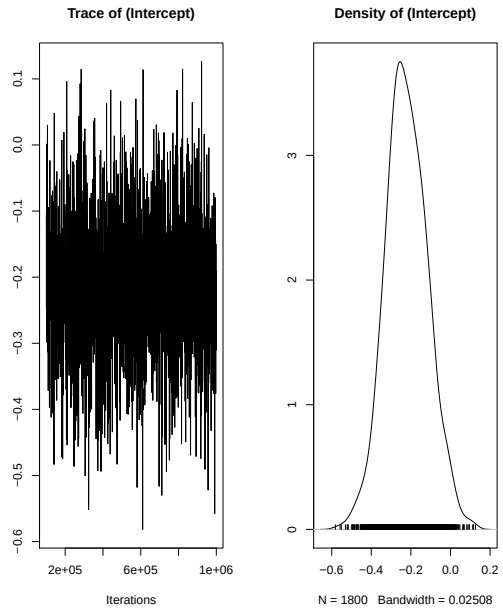

Figure C.1: Trace plot for the effect size.

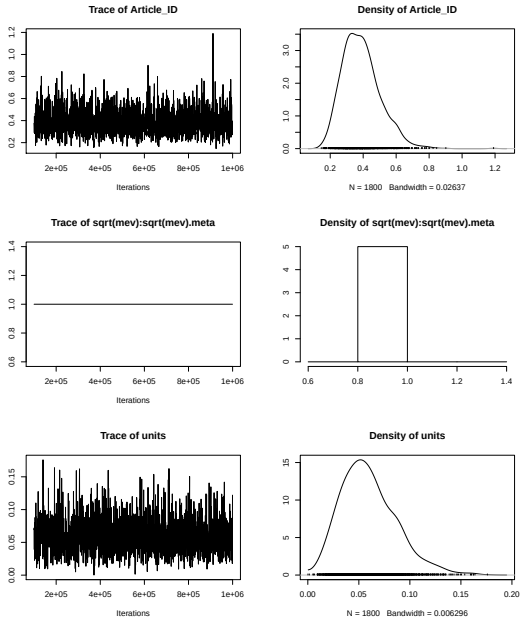

Figure C.2: Trace plot for the variance.

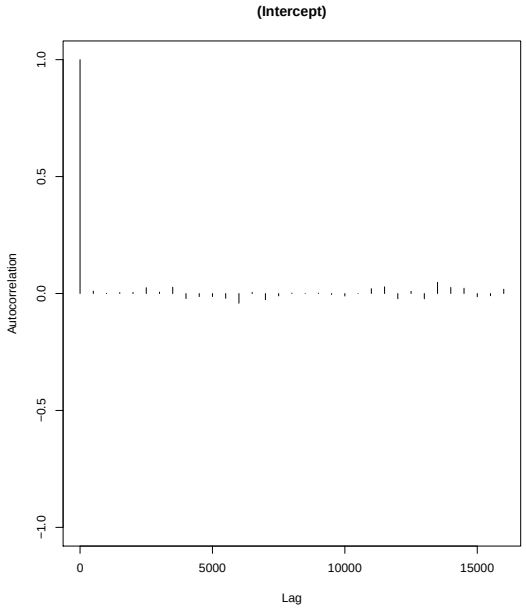

Figure C.3: Autocorrelation of the iterations for the effect size

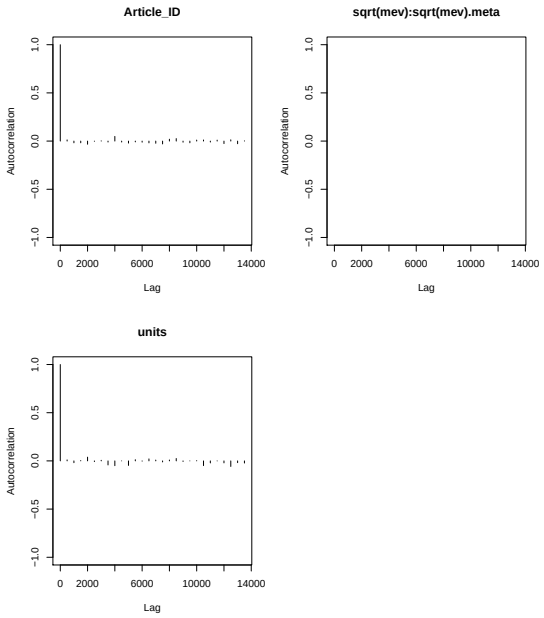

Figure C.4: Autocorrelation of the iterations for the variance.

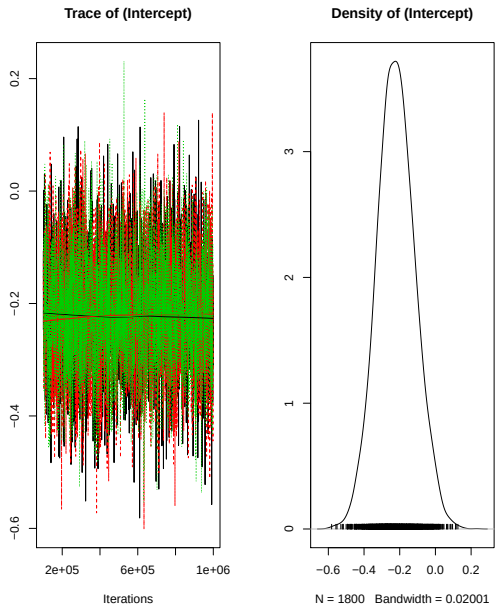

Figure C.5: Trace plot of 3 chains showing similar mixing in all of them.

Appendix D

R code used within Chapter 4

R code used to analyse the data in chapter 4.

```
#This is the journal of runs for the analysis of the meta-analysis data of the
#binary dispersal SMD data. Here I will keep the recorded runs of which
#scientific questions are being asked.


setwd("M:/PhD files/R code/Meta-analysis/SMD data (study 1858 removed)")

#Get the packages required.
library("MCMCglmm")           # for meta-analysis
library("dplyr")              # for data manipulation
library("readxl")             # for reading excel spreadsheets.
library("metafor")            # for additoinal diagnostics.
library("coda")

#Read in the data.
#SMD.data <- read.csv("C:/Users/Vinnie/Dropbox/PhD files/Chapter 3 - Meta-analysis/Meta-analysis 280818/SMD da
SMD.data <- read.csv("M:/PhD files/R code/Meta-analysis/SMD data/SMD_data_Mdrive_Var.csv",
na = "NA")

#remove unwanted rows at the bottom.
SMD.data <- SMD.data[1:225,]
SMD.data <- SMD.data[-c(23:38),]            #Spurious study.
toDelete <- seq(0, nrow(SMD.data), 2)
SMD.data <- SMD.data[-c(toDelete) ,]
#WE now need to remove some nuisance factors from the data.
```

```
for (i in 1:length(SMD.data)){
if (is.factor(SMD.data[,i]) == TRUE){
SMD.data[,i] <- as.character(SMD.data[,i])
SMD.data[,i] <- as.factor(SMD.data[,i])

} else {
SMD.data[,i] <- SMD.data[,i]
}
}
View(SMD.data)


SMD.data.climate <- SMD.data[-which(SMD.data$KG_climate_higher == "N/A"),]
SMD.data.climate$KG_climate_higher <- factor(SMD.data.climate$KG_climate_higher)

SMD.data.lineage <- SMD.data[-which(SMD.data$Environ_lin == "N/A"),]
SMD.data.lineage$Environ_lin <- factor(SMD.data.lineage$Environ_lin)

#The first step is to assess publication bias within the data, which can be done
#using Regger's regression test for asymmetery and examining funnel plots.
#Store the measurement error variance (mev) as a separate vector for entry into the model.

SMD.mev <- SMD.data$SMD_Var_g
SMD.mev.climate <- SMD.data.climate$SMD_Var_g
SMD.mev.lineage <- SMD.data.lineage$SMD_Var_g


prior2 <- list(R = list(V = 1e-10, nu= -1),
G = list(G1 = list(V = 1e-10, nu = -1)))


ES.model.ran_artID <- MCMCglmm(SMD_g ~ 1,  random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.model.ran_artID.1 <- MCMCglmm(SMD_g ~ 1, random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500, prior = prior2 , verbose = FALSE)
ES.model.ran_artID.2 <- MCMCglmm(SMD_g ~ 1, random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500, prior = prior2 , verbose = FALSE)

ES.model.ran_artID.chains <- mcmc.list(ES.model.ran_artID$Sol,ES.model.ran_artID.1$Sol,
ES.model.ran_artID.2$Sol)

plot(ES.model.ran_artID$Sol)
plot(ES.model.ran_artID$VCV)

autocorr.plot(ES.model.ran_artID$Sol)
```

```
autocorr.plot(ES.model.ran_artID$VCV)
effectiveSize(ES.model.ran_artID$Sol)
effectiveSize(ES.model.ran_artID$VCV)
heidel.diag(ES.model.ran_artID$VCV)

summary(ES.model.ran_artID)

#Gelman diagnostics

plot(ES.model.ran_artID.chains)
gelman.diag(ES.model.ran_artID.chains)

ES.Disp_type.model.ran_artID <- MCMCglmm(SMD_g ~ Disp_type - 1, random = ~Article_ID,
data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Disp_type - 1, random = ~Article_ID,
data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Disp_type - 1,  random = ~Article_ID,
data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Disp_type.model.ran_artID.chains <- mcmc.list(ES.Disp_type.model.ran_artID$Sol,
ES.Disp_type.model.ran_artID.1$Sol,
ES.Disp_type.model.ran_artID.2$Sol)

plot(ES.Disp_type.model.ran_artID$Sol)
plot(ES.Disp_type.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Disp_type.model.ran_artID$Sol)
autocorr.plot(ES.Disp_type.model.ran_artID$VCV)
effectiveSize(ES.Disp_type.model.ran_artID$Sol)
effectiveSize(ES.Disp_type.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.Disp_type.model.ran_artID.c$VCV)

#Autocorrelation seems fine.

summary(ES.Disp_type.model.ran_artID)

#Gelman diagnostics

plot(ES.Disp_type.model.ran_artID.chains)
gelman.diag(ES.Disp_type.model.ran_artID.chains)

#We can now do the same for non-dispersal traits.
```

```
ES.fitness_type.model.ran_artID <- MCMCglmm(SMD_g ~ Fitness_type - 1,  random = ~Article_ID,
 data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Fitness_type - 1,  random = ~Article_ID,
 data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Fitness_type - 1,  random = ~Article_ID,
 data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.fitness_type.model.ran_artID.chains <- mcmc.list(ES.fitness_type.model.ran_artID$Sol,
ES.fitness_type.model.ran_artID.1$Sol,
ES.fitness_type.model.ran_artID.2$Sol)

plot(ES.fitness_type.model.ran_artID$Sol)
plot(ES.fitness_type.model.ran_artID$VCV)



autocorr.plot(ES.fitness_type.model.ran_artID$Sol)
autocorr.plot(ES.fitness_type.model.ran_artID$VCV)
effectiveSize(ES.fitness_type.model.ran_artID$Sol)
effectiveSize(ES.fitness_type.model.ran_artID$VCV)

heidel.diag(ES.fitness_type.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.fitness_type.model.ran_artID)

#Gelman diagnostics

plot(ES.fitness_type.model.ran_artID.chains)
gelman.diag(ES.fitness_type.model.ran_artID.chains)

ES.species_class.model.ran_artID <- MCMCglmm(SMD_g ~ Species_class - 1,
random = ~Article_ID, data = SMD.data, mev = SMD.mev, nitt = 1e6,
burnin = 1e5, thin = 500, prior = prior2 , verbose = FALSE)
ES.species_class.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Species_class - 1,
random = ~Article_ID, data = SMD.data, mev = SMD.mev, nitt = 1e6,
burnin = 1e5, thin = 500,prior = prior2 , verbose = FALSE)
ES.species_class.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Species_class - 1,
 random = ~Article_ID, data = SMD.data, mev = SMD.mev, nitt = 1e6,
  burnin = 1e5, thin = 500, prior = prior2 , verbose = FALSE)

ES.species_class.model.ran_artID.chains <- mcmc.list(ES.species_class.model.ran_artID$Sol,
ES.species_class.model.ran_artID.1$Sol,
ES.species_class.model.ran_artID.2$Sol)
```

```
#Check convergence

plot(ES.species_class.model.ran_artID$Sol)

plot(ES.species_class.model.ran_artID$VCV)

#It seems to have mixed well.

autocorr.plot(ES.species_class.model.ran_artID$Sol)
autocorr.plot(ES.species_class.model.ran_artID$VCV)
effectiveSize(ES.species_class.model.ran_artID$Sol)
effectiveSize(ES.species_class.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.species_class.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.species_class.model.ran_artID)

#gelman diagnostics

plot(ES.species_class.model.ran_artID.chains)
gelman.diag(ES.species_class.model.ran_artID.chains)

ES.gender.model.ran_artID <- MCMCglmm(SMD_g ~ Gender - 1,  random = ~Article_ID,
data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Gender - 1,  random = ~Article_ID,
 data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Gender - 1,  random = ~Article_ID,
data = SMD.data, mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.gender.model.ran_artID.chains <- mcmc.list(ES.gender.model.ran_artID$Sol,
ES.gender.model.ran_artID.1$Sol,
ES.gender.model.ran_artID.2$Sol)

#Check convergence

plot(ES.gender.model.ran_artID$Sol)

plot(ES.gender.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.gender.model.ran_artID$Sol)
autocorr.plot(ES.gender.model.ran_artID$VCV)
```

```
effectiveSize(ES.gender.model.ran_artID$Sol)
effectiveSize(ES.gender.model.ran_artID$VCV)

heidel.diag(ES.gender.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.gender.model.ran_artID)

#gelman diagnostics

plot(ES.gender.model.ran_artID.chains)
gelman.diag(ES.gender.model.ran_artID.chains)

#Now climate data
ES.climate.model.ran_artID <- MCMCglmm(SMD_g ~ KG_climate_higher - 1,  random = ~Article_ID, data = S
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.model.ran_artID.1 <- MCMCglmm(SMD_g ~ KG_climate_higher - 1,  random = ~Article_ID, data =
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.model.ran_artID.2 <- MCMCglmm(SMD_g ~ KG_climate_higher - 1,  random = ~Article_ID, data =
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.climate.model.ran_artID.chains <- mcmc.list(ES.climate.model.ran_artID$Sol,
ES.climate.model.ran_artID.1$Sol,
ES.climate.model.ran_artID.2$Sol)

#Check convergence

plot(ES.climate.model.ran_artID$Sol)

plot(ES.climate.model.ran_artID$VCV)

autocorr.plot(ES.climate.model.ran_artID$Sol)
autocorr.plot(ES.climate.model.ran_artID$VCV)
effectiveSize(ES.climate.model.ran_artID$Sol)
effectiveSize(ES.climate.model.ran_artID$VCV)

heidel.diag(ES.climate.model.ran_artID$VCV)

#Passed both tests here.

summary(ES.climate.model.ran_artID)

#gelman diagnostics
```

```
plot(ES.climate.model.ran_artID.chains)
gelman.diag(ES.climate.model.ran_artID.chains)

#Now climate data
ES.environ_lin.model.ran_artID <- MCMCglmm(SMD_g ~ Environ_lin - 1,  random = ~Article_ID, data = SMD.data.lin
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.environ_lin.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Environ_lin - 1,  random = ~Article_ID, data = SMD.data.l
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.environ_lin.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Environ_lin - 1,  random = ~Article_ID, data = SMD.data.l
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.environ_lin.model.ran_artID.chains <- mcmc.list(ES.environ_lin.model.ran_artID$Sol,
ES.environ_lin.model.ran_artID.1$Sol,
ES.environ_lin.model.ran_artID.2$Sol)

#Check convergence

plot(ES.environ_lin.model.ran_artID$Sol)

plot(ES.environ_lin.model.ran_artID$VCV)

#It seems to have mixed well.

autocorr.plot(ES.environ_lin.model.ran_artID$Sol)
autocorr.plot(ES.environ_lin.model.ran_artID$VCV)
effectiveSize(ES.environ_lin.model.ran_artID$Sol)
effectiveSize(ES.environ_lin.model.ran_artID$VCV)

heidel.diag(ES.environ_lin.model.ran_artID$VCV)

#Passed both tests here.

summary(ES.environ_lin.model.ran_artID)

#gelman diagnostics

plot(ES.environ_lin.model.ran_artID.chains)
gelman.diag(ES.environ_lin.model.ran_artID.chains)

#Now climate data
ES.environ_lin.model.int.ran_artID <- MCMCglmm(SMD_g ~ Environ_lin ,  random = ~Article_ID, data = SMD.data.li
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.environ_lin.model.int.ran_artID.1 <- MCMCglmm(SMD_g ~ Environ_lin ,  random = ~Article_ID, data = SMD.data.
```

```
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.environ_lin.model.int.ran_artID.2 <- MCMCglmm(SMD_g ~ Environ_lin ,  random = ~Article_ID, data =
mev = SMD.mev.lineage, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.environ_lin.model.int.ran_artID.chains <- mcmc.list(ES.environ_lin.model.int.ran_artID$Sol,
ES.environ_lin.model.int.ran_artID.1$Sol,
ES.environ_lin.model.int.ran_artID.2$Sol)

#Check convergence

plot(ES.environ_lin.model.int.ran_artID$Sol)

plot(ES.environ_lin.model.int.ran_artID$VCV)

#It seems to have mixed well.

autocorr.plot(ES.environ_lin.model.int.ran_artID$Sol)
autocorr.plot(ES.environ_lin.model.int.ran_artID$VCV)
effectiveSize(ES.environ_lin.model.int.ran_artID$Sol)
effectiveSize(ES.environ_lin.model.int.ran_artID$VCV)

heidel.diag(ES.environ_lin.model.int.ran_artID$VCV)

#Passed both tests here.

summary(ES.environ_lin.model.int.ran_artID)

#gelman diagnostics

plot(ES.environ_lin.model.int.ran_artID.chains)
gelman.diag(ES.environ_lin.model.int.ran_artID.chains)



#We should control for species

#We should now think about running these 2 again and centering them using the code provided.
# #This function should help this
centre2 <- function(nlev){
c1<-contr.treatment(nlev)
c2<-matrix(rep(1/nlev, nlev*(nlev-1)), ncol=nlev-1)
c3<-c1-c2
return(c3)
}

#We do this first for dispersal type, then for Fitness_type.
```

```
SMD.data$Disp_type.c <- SMD.data$Disp_type  # For dispersal type.
contrasts(SMD.data$Disp_type.c)<-centre2(2)

SMD.data$Gender.c <- SMD.data$Gender  # For dispersal type.
contrasts(SMD.data$Gender.c)<-centre2(3)

SMD.data$Fitness_type.c <- SMD.data$Fitness_type  # For Fitness_type
contrasts(SMD.data$Fitness_type.c)<-centre2(3)

SMD.data$Species_class.c <- SMD.data$Species_class  # For Species class
contrasts(SMD.data$Species_class.c)<-centre2(4)

SMD.data.climate$KG_climate_higher.c <- SMD.data.climate$KG_climate_higher  # For Species class
contrasts(SMD.data.climate$KG_climate_higher.c)<-centre2(5)



SMD.data.climate$Disp_type.c <- SMD.data.climate$Disp_type  # For dispersal type.
contrasts(SMD.data.climate$Disp_type.c)<-centre2(2)

SMD.data.climate$Species_class.c <- SMD.data.climate$Species_class  # For Species class
contrasts(SMD.data.climate$Species_class.c)<-centre2(4)

SMD.data.climate$Fitness_type.c <- SMD.data.climate$Fitness_type  # For Fitness_type
contrasts(SMD.data.climate$Fitness_type.c)<-centre2(3)

SMD.data.climate$Gender.c <- SMD.data.climate$Gender  # For dispersal type.
contrasts(SMD.data.climate$Gender.c)<-centre2(3)

ES.Disp_type.Species_class.model.ran_artID <- MCMCglmm(SMD_g ~ Disp_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.Species_class.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Disp_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.Species_class.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Disp_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Disp_type.Species_class.model.ran_artID.chains <- mcmc.list(ES.Disp_type.Species_class.model.ran_artID$Sol,
ES.Disp_type.Species_class.model.ran_artID.1$Sol,
ES.Disp_type.Species_class.model.ran_artID.2$Sol)

plot(ES.Disp_type.Species_class.model.ran_artID$Sol)
```

```
plot(ES.Disp_type.Species_class.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Disp_type.Species_class.model.ran_artID$Sol)
autocorr.plot(ES.Disp_type.Species_class.model.ran_artID$VCV)
effectiveSize(ES.Disp_type.Species_class.model.ran_artID$Sol)
effectiveSize(ES.Disp_type.Species_class.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.Disp_type.Species_class.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.Disp_type.Species_class.model.ran_artID)

#Gelman diagnostics

plot(ES.Disp_type.Species_class.model.ran_artID.chains)
gelman.diag(ES.Disp_type.Species_class.model.ran_artID.chains)



ES.Species_class.Disp_type.model.ran_artID <- MCMCglmm(SMD_g ~ Species_class + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Species_class.Disp_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Species_class + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Species_class.Disp_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Species_class + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Species_class.Disp_type.model.ran_artID.chains <- mcmc.list(ES.Species_class.Disp_type.model.ran_a
ES.Species_class.Disp_type.model.ran_artID.1$Sol,
ES.Species_class.Disp_type.model.ran_artID.2$Sol)

plot(ES.Species_class.Disp_type.model.ran_artID$Sol)
plot(ES.Species_class.Disp_type.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Species_class.Disp_type.model.ran_artID$Sol)
autocorr.plot(ES.Species_class.Disp_type.model.ran_artID$VCV)
effectiveSize(ES.Species_class.Disp_type.model.ran_artID$Sol)
effectiveSize(ES.Species_class.Disp_type.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.Species_class.Disp_type.model.ran_artID$VCV)
```

```
#Autocorrelation seems fine.

summary(ES.Species_class.Disp_type.model.ran_artID)

#Gelman diagnostics

plot(ES.Species_class.Disp_type.model.ran_artID.chains)
gelman.diag(ES.Species_class.Disp_type.model.ran_artID.chains)

#####

ES.gender.Disp_type.model.ran_artID <- MCMCglmm(SMD_g ~ Gender + Disp_type.c - 1,  random = ~Article_ID, data
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.Disp_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Gender + Disp_type.c - 1,  random = ~Article_ID, dat
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.Disp_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Gender + Disp_type.c - 1,  random = ~Article_ID, dat
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.gender.Disp_type.model.ran_artID.chains <- mcmc.list(ES.gender.Disp_type.model.ran_artID$Sol,
ES.gender.Disp_type.model.ran_artID.1$Sol,
ES.gender.Disp_type.model.ran_artID.2$Sol)

#Check convergence

plot(ES.gender.Disp_type.model.ran_artID$Sol)

plot(ES.gender.Disp_type.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.gender.Disp_type.model.ran_artID$Sol)
autocorr.plot(ES.gender.Disp_type.model.ran_artID$VCV)
effectiveSize(ES.gender.Disp_type.model.ran_artID$Sol)
effectiveSize(ES.gender.Disp_type.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.gender.Disp_type.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.gender.Disp_type.model.ran_artID)

#gelman diagnostics

plot(ES.gender.Disp_type.model.ran_artID.chains)
gelman.diag(ES.gender.Disp_type.model.ran_artID.chains)
```

```
#####

ES.Disp_type.gender.model.ran_artID <- MCMCglmm(SMD_g ~ Disp_type + Gender.c - 1,  random = ~Article_
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.gender.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Disp_type + Gender.c - 1,  random = ~Articl
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.gender.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Disp_type + Gender.c - 1,  random = ~Articl
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Disp_type.gender.model.ran_artID.chains <- mcmc.list(ES.Disp_type.gender.model.ran_artID$Sol,
ES.Disp_type.gender.model.ran_artID.1$Sol,
ES.Disp_type.gender.model.ran_artID.2$Sol)

#Check convergence

plot(ES.Disp_type.gender.model.ran_artID$Sol)

plot(ES.Disp_type.gender.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Disp_type.gender.model.ran_artID$Sol)
autocorr.plot(ES.Disp_type.gender.model.ran_artID$VCV)
effectiveSize(ES.Disp_type.gender.model.ran_artID$Sol)
effectiveSize(ES.Disp_type.gender.model.ran_artID$VCV)

heidel.diag(ES.Disp_type.gender.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.Disp_type.gender.model.ran_artID)

#gelman diagnostics

plot(ES.Disp_type.gender.model.ran_artID.chains)
gelman.diag(ES.Disp_type.gender.model.ran_artID.chains)

#######################


ES.climate.Species_class.model.ran_artID <- MCMCglmm(SMD_g ~ KG_climate_higher + Species_class.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
```

```
ES.climate.Species_class.model.ran_artID.1 <- MCMCglmm(SMD_g ~ KG_climate_higher + Species_class.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.Species_class.model.ran_artID.2 <- MCMCglmm(SMD_g ~ KG_climate_higher + Species_class.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.climate.Species_class.model.ran_artID.chains <- mcmc.list(ES.climate.Species_class.model.ran_artID$Sol,
ES.climate.Species_class.model.ran_artID.1$Sol,
ES.climate.Species_class.model.ran_artID.2$Sol)

#Check convergence

plot(ES.climate.Species_class.model.ran_artID$Sol)

plot(ES.climate.Species_class.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.climate.Species_class.model.ran_artID$Sol)
autocorr.plot(ES.climate.Species_class.model.ran_artID$VCV)
effectiveSize(ES.climate.Species_class.model.ran_artID$Sol)
effectiveSize(ES.climate.Species_class.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.climate.Species_class.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.climate.Species_class.model.ran_artID)

#gelman diagnostics

plot(ES.climate.Species_class.model.ran_artID.chains)
gelman.diag(ES.climate.Species_class.model.ran_artID.chains)

##########################


ES.Species_class.climate.model.ran_artID <- MCMCglmm(SMD_g ~ Species_class + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Species_class.climate.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Species_class + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
```

```
ES.Species_class.climate.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Species_class + KG_climate_higher.c -
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Species_class.climate.model.ran_artID.chains <- mcmc.list(ES.Species_class.climate.model.ran_artID
ES.Species_class.climate.model.ran_artID.1$Sol,
ES.Species_class.climate.model.ran_artID.2$Sol)

#Check convergence

plot(ES.Species_class.climate.model.ran_artID$Sol)

plot(ES.Species_class.climate.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Species_class.climate.model.ran_artID$Sol)
autocorr.plot(ES.Species_class.climate.model.ran_artID$VCV)
effectiveSize(ES.Species_class.climate.model.ran_artID$Sol)
effectiveSize(ES.Species_class.climate.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.Species_class.climate.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.Species_class.climate.model.ran_artID)

#gelman diagnostics

plot(ES.Species_class.climate.model.ran_artID.chains)
gelman.diag(ES.Species_class.climate.model.ran_artID.chains)

############################


ES.Disp_type.climate.model.ran_artID <- MCMCglmm(SMD_g ~ Disp_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.climate.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Disp_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.climate.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Disp_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
```

```
ES.Disp_type.climate.model.ran_artID.chains <- mcmc.list(ES.Disp_type.climate.model.ran_artID$Sol,
ES.Disp_type.climate.model.ran_artID.1$Sol,
ES.Disp_type.climate.model.ran_artID.2$Sol)

#Check convergence

plot(ES.Disp_type.climate.model.ran_artID$Sol)

plot(ES.Disp_type.climate.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Disp_type.climate.model.ran_artID$Sol)
autocorr.plot(ES.Disp_type.climate.model.ran_artID$VCV)
effectiveSize(ES.Disp_type.climate.model.ran_artID$Sol)
effectiveSize(ES.Disp_type.climate.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.Disp_type.climate.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.Disp_type.climate.model.ran_artID)

#gelman diagnostics

plot(ES.Disp_type.climate.model.ran_artID.chains)
gelman.diag(ES.Disp_type.climate.model.ran_artID.chains)

####################################

ES.climate.Disp_type.model.ran_artID <- MCMCglmm(SMD_g ~ KG_climate_higher + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.Disp_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~ KG_climate_higher + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.Disp_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~ KG_climate_higher + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.climate.Disp_type.model.ran_artID.chains <- mcmc.list(ES.climate.Disp_type.model.ran_artID$Sol,
ES.climate.Disp_type.model.ran_artID.1$Sol,
ES.climate.Disp_type.model.ran_artID.2$Sol)

#Check convergence
```

```
plot(ES.climate.Disp_type.model.ran_artID$Sol)

plot(ES.climate.Disp_type.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.climate.Disp_type.model.ran_artID$Sol)
autocorr.plot(ES.climate.Disp_type.model.ran_artID$VCV)
effectiveSize(ES.climate.Disp_type.model.ran_artID$Sol)
effectiveSize(ES.climate.Disp_type.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.climate.Disp_type.model.ran_artID$VCV)
#Passed both tests here.

#gelman diagnostics

plot(ES.climate.Disp_type.model.ran_artID.chains)
gelman.diag(ES.climate.Disp_type.model.ran_artID.chains)

################################################################

#Now we run the centred versions for dispersal type.

ES.Disp_type.fitness.model.ran_artID <- MCMCglmm(SMD_g ~ Disp_type + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.fitness.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Disp_type + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Disp_type.fitness.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Disp_type + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Disp_type.fitness.model.ran_artID.chains <- mcmc.list(ES.Disp_type.fitness.model.ran_artID$Sol,
ES.Disp_type.fitness.model.ran_artID.1$Sol,
ES.Disp_type.fitness.model.ran_artID.2$Sol)

plot(ES.Disp_type.fitness.model.ran_artID$Sol)
plot(ES.Disp_type.fitness.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Disp_type.fitness.model.ran_artID$Sol)
autocorr.plot(ES.Disp_type.fitness.model.ran_artID$VCV)
effectiveSize(ES.Disp_type.fitness.model.ran_artID$Sol)
effectiveSize(ES.Disp_type.fitness.model.ran_artID$VCV)
```

```
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.Disp_type.fitness.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.Disp_type.fitness.model.ran_artID)

#Gelman diagnostics

plot(ES.Disp_type.fitness.model.ran_artID.chains)
gelman.diag(ES.Disp_type.fitness.model.ran_artID.chains)

######################################################################

#Now we run the centred versions for dispersal type.

ES.fitness.Disp_type.model.ran_artID <- MCMCglmm(SMD_g ~ Fitness_type + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.Disp_type.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Fitness_type + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.Disp_type.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Fitness_type + Disp_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.fitness.Disp_type.model.ran_artID.chains <- mcmc.list(ES.fitness.Disp_type.model.ran_artID$Sol,
ES.fitness.Disp_type.model.ran_artID.1$Sol,
ES.fitness.Disp_type.model.ran_artID.2$Sol)


plot(ES.fitness.Disp_type.model.ran_artID$Sol)
plot(ES.fitness.Disp_type.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.fitness.Disp_type.model.ran_artID$Sol)
autocorr.plot(ES.fitness.Disp_type.model.ran_artID$VCV)
effectiveSize(ES.fitness.Disp_type.model.ran_artID$Sol)
effectiveSize(ES.fitness.Disp_type.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.fitness.Disp_type.model.ran_artID$VCV)

#Autocorrelation seems fine.
```

```
summary(ES.fitness.Disp_type.model.ran_artID)

#Gelman diagnostics

plot(ES.fitness.Disp_type.model.ran_artID.chains)
gelman.diag(ES.fitness.Disp_type.model.ran_artID.chains)

#####################################################################

ES.fitness.Species_class.model.ran_artID <- MCMCglmm(SMD_g ~ Fitness_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.Species_class.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Fitness_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.Species_class.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Fitness_type + Species_class.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.fitness.Species_class.model.ran_artID.chains <- mcmc.list(ES.fitness.Species_class.model.ran_artID
ES.fitness.Species_class.model.ran_artID.1$Sol,
ES.fitness.Species_class.model.ran_artID.2$Sol)

plot(ES.fitness.Species_class.model.ran_artID$Sol)
plot(ES.fitness.Species_class.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.fitness.Species_class.model.ran_artID$Sol)
autocorr.plot(ES.fitness.Species_class.model.ran_artID$VCV)
effectiveSize(ES.fitness.Species_class.model.ran_artID$Sol)
effectiveSize(ES.fitness.Species_class.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.fitness.Species_class.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.fitness.Species_class.model.ran_artID)

#Gelman diagnostics

plot(ES.fitness.Species_class.model.ran_artID.chains)
gelman.diag(ES.fitness.Species_class.model.ran_artID.chains)

#####################################################################
```

```
#Now we run the centred versions for dispersal type.

ES.Species_class.fitness.model.ran_artID <- MCMCglmm(SMD_g ~ Species_class + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Species_class.fitness.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Species_class + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.Species_class.fitness.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Species_class + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.Species_class.fitness.model.ran_artID.chains <- mcmc.list(ES.Species_class.fitness.model.ran_artID$Sol,
ES.Species_class.fitness.model.ran_artID.1$Sol,
ES.Species_class.fitness.model.ran_artID.2$Sol)


plot(ES.Species_class.fitness.model.ran_artID$Sol)
plot(ES.Species_class.fitness.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.Species_class.fitness.model.ran_artID$Sol)
autocorr.plot(ES.Species_class.fitness.model.ran_artID$VCV)
effectiveSize(ES.Species_class.fitness.model.ran_artID$Sol)
effectiveSize(ES.Species_class.fitness.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.Species_class.fitness.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.Species_class.fitness.model.ran_artID)

#Gelman diagnostics

plot(ES.Species_class.fitness.model.ran_artID.chains)
gelman.diag(ES.Species_class.fitness.model.ran_artID.chains)

#####################################################################

#Now we run the centred versions for dispersal type.

ES.fitness.gender.model.ran_artID <- MCMCglmm(SMD_g ~ Fitness_type + Gender.c - 1,  random = ~Article_ID, data
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
```

```
prior = prior2 , verbose = FALSE)
ES.fitness.gender.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Fitness_type + Gender.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.gender.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Fitness_type + Gender.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.fitness.gender.model.ran_artID.chains <- mcmc.list(ES.fitness.gender.model.ran_artID$Sol,
ES.fitness.gender.model.ran_artID.1$Sol,
ES.fitness.gender.model.ran_artID.2$Sol)

plot(ES.fitness.gender.model.ran_artID$Sol)
plot(ES.fitness.gender.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.fitness.gender.model.ran_artID$Sol)
autocorr.plot(ES.fitness.gender.model.ran_artID$VCV)
effectiveSize(ES.fitness.gender.model.ran_artID$Sol)
effectiveSize(ES.fitness.gender.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.fitness.gender.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.fitness.gender.model.ran_artID)

#Gelman diagnostics

plot(ES.fitness.gender.model.ran_artID.chains)
gelman.diag(ES.fitness.gender.model.ran_artID.chains)

######################################################################

#Now we run the centred versions for dispersal type.

ES.gender.fitness.model.ran_artID <- MCMCglmm(SMD_g ~ Gender + Fitness_type.c - 1,  random = ~Article
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.fitness.model.ran_artID.1 <- MCMCglmm(SMD_g ~  Gender + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.fitness.model.ran_artID.2 <- MCMCglmm(SMD_g ~  Gender + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data,
```

```
mev = SMD.mev, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.gender.fitness.model.ran_artID.chains <- mcmc.list(ES.gender.fitness.model.ran_artID$Sol,
ES.gender.fitness.model.ran_artID.1$Sol,
ES.gender.fitness.model.ran_artID.2$Sol)

#I received  ino warnings for this model, poerhaps that is what was wrong before.

plot(ES.gender.fitness.model.ran_artID$Sol)
plot(ES.gender.fitness.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.gender.fitness.model.ran_artID$Sol)
autocorr.plot(ES.gender.fitness.model.ran_artID$VCV)
effectiveSize(ES.gender.fitness.model.ran_artID$Sol)
effectiveSize(ES.gender.fitness.model.ran_artID$VCV)
#effective sample sizes were 900, so pretty good.
heidel.diag(ES.gender.fitness.model.ran_artID$VCV)

#Autocorrelation seems fine.

summary(ES.gender.fitness.model.ran_artID)

#Gelman diagnostics

plot(ES.gender.fitness.model.ran_artID.chains)
gelman.diag(ES.gender.fitness.model.ran_artID.chains)

################################################################

ES.fitness.climate.model.ran_artID <- MCMCglmm(SMD_g ~ Fitness_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.climate.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Fitness_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.fitness.climate.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Fitness_type + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.fitness.climate.model.ran_artID.chains <- mcmc.list(ES.fitness.climate.model.ran_artID$Sol,
ES.fitness.climate.model.ran_artID.1$Sol,
ES.fitness.climate.model.ran_artID.2$Sol)
```

```
#Check convergence

plot(ES.fitness.climate.model.ran_artID$Sol)

plot(ES.fitness.climate.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.fitness.climate.model.ran_artID$Sol)
autocorr.plot(ES.fitness.climate.model.ran_artID$VCV)
effectiveSize(ES.fitness.climate.model.ran_artID$Sol)
effectiveSize(ES.fitness.climate.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.fitness.climate.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.fitness.climate.model.ran_artID)

#gelman diagnostics

plot(ES.fitness.climate.model.ran_artID.chains)
gelman.diag(ES.fitness.climate.model.ran_artID.chains)

#############################################################

ES.climate.fitness.model.ran_artID <- MCMCglmm(SMD_g ~ KG_climate_higher + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.fitness.model.ran_artID.1 <- MCMCglmm(SMD_g ~ KG_climate_higher + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.fitness.model.ran_artID.2 <- MCMCglmm(SMD_g ~ KG_climate_higher + Fitness_type.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.climate.fitness.model.ran_artID.chains <- mcmc.list(ES.climate.fitness.model.ran_artID$Sol,
ES.climate.fitness.model.ran_artID.1$Sol,
ES.climate.fitness.model.ran_artID.2$Sol)

#Check convergence

plot(ES.climate.fitness.model.ran_artID$Sol)

plot(ES.climate.fitness.model.ran_artID$VCV)
```

```
#It seems to have mixed well.
autocorr.plot(ES.climate.fitness.model.ran_artID$Sol)
autocorr.plot(ES.climate.fitness.model.ran_artID$VCV)
effectiveSize(ES.climate.fitness.model.ran_artID$Sol)
effectiveSize(ES.climate.fitness.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.climate.fitness.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.climate.fitness.model.ran_artID)

#gelman diagnostics

plot(ES.climate.fitness.model.ran_artID.chains)
gelman.diag(ES.climate.fitness.model.ran_artID.chains)

##############################################################

ES.climate.gender.model.ran_artID <- MCMCglmm(SMD_g ~ KG_climate_higher + Gender.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.gender.model.ran_artID.1 <- MCMCglmm(SMD_g ~ KG_climate_higher + Gender.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.climate.gender.model.ran_artID.2 <- MCMCglmm(SMD_g ~ KG_climate_higher + Gender.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.climate.gender.model.ran_artID.chains <- mcmc.list(ES.climate.gender.model.ran_artID$Sol,
ES.climate.gender.model.ran_artID.1$Sol,
ES.climate.gender.model.ran_artID.2$Sol)

#Check convergence

plot(ES.climate.gender.model.ran_artID$Sol)

plot(ES.climate.gender.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.climate.gender.model.ran_artID$Sol)
autocorr.plot(ES.climate.gender.model.ran_artID$VCV)
effectiveSize(ES.climate.gender.model.ran_artID$Sol)
effectiveSize(ES.climate.gender.model.ran_artID$VCV)
```

```
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.climate.gender.model.ran_artID$VCV)
#Passed both tests here.

summary(ES.climate.gender.model.ran_artID)
#DIC here is -8.744418. We have significance of females and plants, but the plants are from one artic

#gelman diagnostics

plot(ES.climate.gender.model.ran_artID.chains)
gelman.diag(ES.climate.gender.model.ran_artID.chains)

#############################################################

ES.gender.climate.model.ran_artID <- MCMCglmm(SMD_g ~ Gender + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.climate.model.ran_artID.1 <- MCMCglmm(SMD_g ~ Gender + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)
ES.gender.climate.model.ran_artID.2 <- MCMCglmm(SMD_g ~ Gender + KG_climate_higher.c - 1,
random = ~Article_ID, data = SMD.data.climate,
mev = SMD.mev.climate, nitt = 1e6, burnin = 1e5, thin = 500,
prior = prior2 , verbose = FALSE)

ES.gender.climate.model.ran_artID.chains <- mcmc.list(ES.gender.climate.model.ran_artID$Sol,
ES.gender.climate.model.ran_artID.1$Sol,
ES.gender.climate.model.ran_artID.2$Sol)

#Check convergence

plot(ES.gender.climate.model.ran_artID$Sol)

plot(ES.gender.climate.model.ran_artID$VCV)

#It seems to have mixed well.
autocorr.plot(ES.gender.climate.model.ran_artID$Sol)
autocorr.plot(ES.gender.climate.model.ran_artID$VCV)
effectiveSize(ES.gender.climate.model.ran_artID$Sol)
effectiveSize(ES.gender.climate.model.ran_artID$VCV)
#The effective sample sizes have dropped. but autocorrelation still looks ok.
heidel.diag(ES.gender.climate.model.ran_artID$VCV)
#Passed both tests here.
```

```
#gelman diagnostics

plot(ES.gender.climate.model.ran_artID.chains)
gelman.diag(ES.gender.climate.model.ran_artID.chains)
```

Bibliography

[1] Drought-hit cape town dreads 'day zero' when taps will run dry. https://uk.reuters.com/article/uk-safrica-drought/drought-hit-cape-town-dreads-day-zero-when-taps-will-run-dry-idUKKCN1G51P Accessed: 2018-09-26.

[2] International cooperation for animal research using space. http://www.orn.mpg.de/ICARUS. Accessed: 2018-09-26.

[3] Ejaz Ahmad, Iftikhar Hussain, and Joe E. Brooks. Losses of stored foods due to rats at grain markets in pakistan. *International Biodeterioration & Biodegradation*, 36(1-2):125–133, 1995.

[4] Kama N. Almasi. A non-native perennial invades a native forest. *Biological Invasions*, 2(3):219–230, 2000.

[5] D.A. Andow, P. M_ Kareiva, Simon A. Levin, and Akira Okubo. Spread of invading organisms. *Landscape Ecology*, 4(2-3):177–188, 1990.

[6] J.H. Appleby and P.F. Credland. Bionomics and polymorphism in callosobruchus subinnotatus (coleoptera: Bruchidae). *Bulletin of Entomological Research*, 91(4):235–244, 2001.

[7] Michel Baguette and Hans Van Dyck. Landscape connectivity and animal behavior: functional grain as a key determinant for dispersal. *Landscape Ecology*, 22(8):1117–1129, 2007.

[8] Kamil A. Barton, Ben L. Phillips, Juan M. Morales, and Justin M.J. Travis. The evolution of an 'intelligent'dispersal strategy: biased, correlated random walks in patchy landscapes. *Oikos*, 118(2):309–319, 2009.

[9] Céline Bellard, Cleo Bertelsmeier, Paul Leadley, Wilfried Thuiller, and Franck Courchamp. Impacts of climate change on the future of biodiversity. *Ecology Letters*, 15(4):365–377, 2012.

[10] Tim M. Blackburn, Petr Pyšek, Sven Bacher, James T. Carlton, Richard P. Duncan, Vojtěch Jarošík, John R.U. Wilson, and David M. Richardson. A proposed unified framework for biological invasions. *Trends in Ecology & Evolution*, 26(7):333–339, 2011.

[11] Patrick J. Bohlen, Stefan Scheu, Cindy M. Hale, Mary Ann McLean, Sonja Migge, Peter M. Groffman, and Dennis Parkinson. Non-native invasive earthworms as agents of change in northern temperate forests. *Frontiers in Ecology and the Environment*, 2(8):427–435, 2004.

[12] Dries Bonte, Hans Van Dyck, James M. Bullock, Aurélie Coulon, Maria Delgado, Melanie Gibbs, Valerie Lehouck, Erik Matthysen, Karin Mustin, Marjo Saastamoinen, Nicolas Schtickzelle, Virginie M. Stevens, Sofie Vandewoestijne, Michel Baguette, Kamil Barton, Tim G. Benton, Audrey Chaput-Bardy, Jean Clobert, Calvin Dytham, Thomas Hovestadt, Christoph M. Meier, Steve C. F. Palmer, Camille Turlure, and Justin M. J. Travis. Costs of dispersal. *Biological Reviews*, 87:290–312, Sep 2011.

[13] Dries Bonte, Martin Lukáč, and Luc Lens. Starvation affects pre-dispersal behaviour of erigone spiders. *Basic and Applied Ecology*, 9(3):308–315,

2008.

[14] Michael Borenstein, Larry V. Hedges, Julian P.T Higgins, and Hannah R. Rothstein. *Introduction to meta-analysis*. John Wiley & Sons, 2011.

[15] Diana E. Bowler and Tim G. Benton. Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, 80(2):205–225, 2005.

[16] Matthew L. Brooks, Carla M. D'antonio, David M. Richardson, James B. Grace, Jon E. Keeley, Joseph M. DiTomaso, Richard J. Hobbs, Mike Pellant, and David Pyke. Effects of invasive alien plants on fire regimes. *AIBS Bulletin*, 54(7):677–688, 2004.

[17] Olivia J. Burton, Ben L. Phillips, and Justin M. J. Travis. Trade-offs and the evolution of life-histories during range expansion. *Ecology Letters*, 13(10):1210–1220, Aug 2010.

[18] Jarrett E. Byrnes, Pamela L. Reynolds, and John J. Stachowicz. Invasions and extinctions reshape coastal marine food webs. *PloS one*, 2(3):e295, 2007.

[19] Sergio A. Cannas, Diana E. Marco, and A. Sergio Páez. Modelling biological invasions: species traits, species interactions, and habitat heterogeneity. *Mathematical Biosciences*, 183(1):93–110, May 2003.

[20] Ted J. Case, Robert D. Holt, Mark A. McPeek, and Timothy H. Keitt. The community context of species' borders: ecological and evolutionary perspectives. *Oikos*, 108(1):28–46, 2005.

[21] I-C. Chen, J. K. Hill, R. Ohlemuller, D. B. Roy, and C. D. Thomas. Rapid range shifts of species associated with high levels of climate warming. *Sci-*

ence, 333(6045):1024–1026, Aug 2011.

[22] Angela Chuang and Christopher R. Peterson. Expanding population edges: theories, traits, and trade-offs. *Global Change Biology*, 22(2):494–512, Jan 2016.

[23] Martin L. Cipollini and Douglas J. Levey. Secondary metabolites of fleshy vertebrate-dispersed fruits: adaptive hypotheses and implications for seed dispersal. *The American Naturalist*, 150(3):346–372, 1997.

[24] James S. Clark. Why trees migrate so fast: Confronting theory with dispersal biology and the paleorecord. *The American Naturalist*, 152:204–224, Aug 1998.

[25] James S. Clark, Chris Fastie, George Hurtt, Stephen T. Jackson, Carter Johnson, George A. King, Mark Lewis, Jason Lynch, Stephen Pacala, Colin Prentice, Eugene W. Schupp, Thompson Webb, and Peter Wyckoff. Reid's paradox of rapid plant migration. *BioScience*, 48:13–24, Jan 1998.

[26] James S. Clark, Miles Silman, Ruth Kern, Eric Macklin, and Janneke HilleRisLambers. Seed dispersal near and far: patterns across temperate and tropical forests. *Ecology*, 80(5):1475–1494, 1999.

[27] Miguel Clavero and Emili García-Berthou. Invasive species are a leading cause of animal extinctions. *Trends in Ecology & Evolution*, 20(3):110, 2005.

[28] Jean Clobert, Michel Baguette, Tim G. Benton, and James M. Bullock, editors. *Dispersal Ecology and Evolution*. Oxford University Press, Sep 2012.

[29] Jean Clobert, Jean-François Le Galliard, Julien Cote, Sandrine Meylan, and Manuel Massot. Informed dispersal, heterogeneity in animal dispersal

syndromes and the dynamics of spatially structured populations. *Ecology Letters*, 12(3):197–209, 2009.

[30] Jacob Cohen. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1):37–46, 1960.

[31] Robert I. Colautti, Sarah A. Bailey, Colin D.A. Van Overdijk, Keri Amundsen, and Hugh J. MacIsaac. Characterised and projected costs of nonindigenous species in canada. *Biological Invasions*, 8(1):45–59, 2006.

[32] Chris Cosner. Challenges in modeling biological invasions and population distributions in a changing climate. *Ecological Complexity*, 20:258–263, 2014.

[33] J. Cote and J. Clobert. Risky dispersal: avoiding kin competition despite uncertainty. *Ecology*, 91(5):1485–1493, 2010.

[34] J. Crank, P. Nicolson, and R. D. Hartree. A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type. *Mathematical Proceedings of the Cambridge Philosophical Society*, 43(1):50–67, Jan 1947.

[35] Peter Crnokrak and Derek A. Roff. The genetic basis of the trade-off between calling and wing morph in males of the cricket gryllus firmus. *Evolution*, 52(4):1111–1118, 1998.

[36] Jim M. Cushing, Sheree Levarge, Nakul Chitnis, and Shandelle M. Henson. Some discrete competition models and the competitive exclusion principle. *Journal of Difference Equations and Applications*, 10(13-15):1139–1151, 2004.

[37] Carla M. D'Antonio, Timothy J. Tunison, and K. Rhonda Loh. Variation

in the impact of exotic grasses on native plant composition in relation to fire across an elevation gradient in hawaii. *Austral Ecology*, 25(5):507–522, Oct 2000.

[38] Peter Daszak, Andrew A. Cunningham, and Alex D. Hyatt. Emerging infectious diseases of wildlife–threats to biodiversity and human health. *Science*, 287(5452):443–449, 2000.

[39] Zoe G. Davies, Robert J. Wilson, Tom M. Brereton, and Chris D. Thomas. The re-expansion and improving status of the silver-spotted skipper butterfly (hesperia comma) in britain: a metapopulation success story. *Biological Conservation*, 124(2):189–198, 2005.

[40] A. Mark Davis, Philip J. Grime, and Ken Thompson. Fluctuating resources in plant communities: a general theory of invasibility. *Journal of Ecology*, 88(3):528–534, Jun 2000.

[41] Matthieu Delefosse, Katrina Povidisa, Dorine Poncet, Erik Kristensen, and Birgit Olesen. Variation in size and chemical composition of seeds from the seagrass zostera marina—ecological implications. *Aquatic Botany*, 131:7–14, 2016.

[42] Jack Dockery, Vivian Hutson, Konstantin Mischaikow, and Mark Pernarowski. The evolution of slow dispersal rates: a reaction diffusion model. *Journal of Mathematical Biology*, 37(1):61–83, 1998.

[43] Jonathan Dubois and Pierre-Olivier Cheptou. Competition/colonization syndrome mediated by early germination in non-dispersing achenes in the heteromorphic species crepis sancta. *Annals of botany*, 110(6):1245–1251, 2012.

[44] Ruth E. Dunn, Craig R. White, and Jonathan A. Green. A model to esti-

mate seabird field metabolic rates. *Biology Letters*, 14(6):20180190, 2018.

[45] Calvin Dytham. Evolved dispersal strategies at range margins. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1661):1407–1413, Apr 2009.

[46] Matthias Egger, George Davey-Smith, and Douglas Altman. *Systematic reviews in health care: meta-analysis in context*. John Wiley & Sons, 2008.

[47] Elizabeth C. Elliott and Stephen J. Cornell. Dispersal polymorphism and the speed of biological invasions. *PLoS ONE*, 7(7):40496, Jul 2012.

[48] Elizabeth C. Elliott and Stephen J. Cornell. Are anomalous invasion speeds robust to demographic stochasticity? *PLoS ONE*, 8(7):67871, Jul 2013.

[49] Charles S. Elton. *The ecology of invasions by animals and plants*. University of Chicago Press, 1958.

[50] Paul R. Epstein, Henry F. Diaz, Scott Elias, Georg Grabherr, Nicholas E. Graham, Willem J.M. Martens, Ellen MosIey-Thompson, and Joel Susskind. Biological and physical signs of climate change: focus on mosquito-borne diseases. *Bulletin of the American Meteorological Society*, 79(3):409–418, 1998.

[51] Laurent Excoffier and Nicolas Ray. Surfing during population expansions promotes genetic revolutions and structuration. *Trends in Ecology & Evolution*, 23(7):347–351, 2008.

[52] Jordi Figuerola, Iris Charalambidou, Luis Santamaria, and Andy J. Green. Internal dispersal of seeds by waterfowl: effect of seed size on gut passage time and germination patterns. *Naturwissenschaften*, 97(6):555–565, 2010.

[53] Ronald Aylmer Fisher. The wave of advance of advantageous genes. *Annals of Eugenics*, 7(4):355–369, June 1937.

[54] David R. French and Justin M.J. Travis. Density-dependent dispersal in host-parasitoid assemblages. *Oikos*, 95(1):125–135, 2001.

[55] Emanuel A. Fronhofer and Florian Altermatt. Eco-evolutionary feedbacks during experimental range expansions. *Nature Communications*, 6(1), Apr 2015.

[56] Kenji Fujisaki. Reproduction and egg diapause of the oriental chinch bug, cavelerius saccharivorus okajima (heteroptera: Lygaeidae), in the subtropical winter season in relation to its wing polymorphism. *Researches on Population Ecology*, 35(2):171–181, 1993.

[57] K.N. Ganeshaiah and Uma R. Shaanker. Seed size optimization in a wind dispersed tree butea monosperma: a trade-off between seedling establishment and pod dispersal efficiency. *Oikos*, pages 3–6, 1991.

[58] Gisela Garcia-Ramos and Mark Kirkpatrick. Genetic models of adaptation and gene flow in peripheral populations. *Evolution*, 51(1):21, Feb 1997.

[59] Andrew Gelman et al. Prior distributions for variance parameters in hierarchical models (comment on article by browne and draper). *Bayesian Analysis*, 1(3):515–534, 2006.

[60] Andrew Gelman, Donald B. Rubin, et al. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7(4):457–472, 1992.

[61] L. Girardin. Non-cooperative fisher-kpp systems: traveling waves and long-time behavior. *Nonlinearity*, 31:108–164, 2018.

[62] Léo Girardin. Non-cooperative fisher–kpp systems: traveling waves and long-time behavior. *Nonlinearity*, 31(1):108–164, Dec 2017.

[63] H. Charles J. Godfray, John R. Beddington, Ian R. Crute, Lawrence Haddad, David Lawrence, James F Muir, Jules Pretty, Sherman Robinson, Sandy M Thomas, and Camilla Toulmin. Food security: the challenge of feeding 9 billion people. *Science*, page 1185383, 2010.

[64] Laura I. González-Guzmán and David W. Mehlman. Developmental stability across the breeding distribution of the scissor-tailed flycatcher (tyrannus forficatus). *Ecology Letters*, 4(5):444–452, 2001.

[65] Paul J. Greenwood. Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, 28(4):1140–1162, 1980.

[66] Patrick A. Guerra. Evaluating the life-history trade-off between dispersal capability and reproduction in wing dimorphic insects: a meta-analysis. *Biological Reviews*, 86(4):813–835, Jan 2011.

[67] Jessica Gurevitch and Dianna K. Padilla. Are invasive species a major cause of extinctions? *Trends in Ecology & Evolution*, 19(9):470–474, 2004.

[68] Yitzchak Gutterman. Strategies of seed dispersal and germination in plants inhabiting deserts. *The Botanical Review*, 60(4):373–425, 1994.

[69] J. D. Hadfield and S. Nakagawa. General quantitative genetic methods for comparative biology: phylogenies, taxonomies and multi-trait models for continuous and categorical characters. *Journal of Evolutionary Biology*, 23(3):494–508, Mar 2010.

[70] Jarrod D. Hadfield. Mcmc methods for multi-response generalized linear mixed models: The MCMCglmm R package. *Journal of Statistical Software*,

33(2):1–22, 2010.

[71] J. B. S. Haldane. The relation between density regulation and natural selection. *Proceedings of the Royal Society B: Biological Sciences*, 145(920):306–308, Jul 1956.

[72] W. D. Hamilton and Robert M. May. Dispersal in stable habitats. *Nature*, 269(5629):578–581, Oct 1977.

[73] Ilkka Hanski. A practical model of metapopulation dynamics. *Journal of Animal Ecology*, pages 151–162, 1994.

[74] Ilkka Hanski, Michael E. Gilpin, and David E. McCauley. *Metapopulation Biology*, volume 454. Elsevier, 1997.

[75] Ilkka Hanski, Marjo Saastamoinen, and Otso Ovaskainen. Dispersal-related life-history trade-offs in a butterfly metapopulation. *Journal of Animal Ecology*, 75:91–100, 2006.

[76] Lennart Hansson. Dispersal and connectivity in metapopulations. In *Metapopulation dynamics: empirical and theoretical investigations*, pages 89–103. Elsevier, 1991.

[77] Deborah R. Hart and Robert H. Gardner. A spatial model for the spread of invading organisms subject to competition. *Journal of Mathematical Biology*, 35(8):935–948, Sep 1997.

[78] Alan Hastings, Kim Cuddington, Kendi F. Davies, Christopher J. Dugaw, Sarah Elmendorf, Amy Freestone, Susan Harrison, Matthew Holland, John Lambrinos, Urmila Malvadkar, Brett A. Melbourne, Kara Moore, Caz Taylor, and Diane Thomson. The spatial spread of invasions: new developments in theory and evidence. *Ecology Letters*, pages 91–101, 2004.

[79] José L. Hierro, Christopher J. Lortie, Diego Villarreal, María E Estanga-Mollica, and Ragan M Callaway. Resistance to centaurea solstitialis invasion from annual and perennial grasses in california and argentina. *Biological Invasions*, 13(10):2249–2259, 2011.

[80] Julian P.T. Higgins and Simon G. Thompson. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, 21(11):1539–1558, 2002.

[81] J.K. Hill, C.D. Thomas, and D.S. Blakeley. Evolution of flight morphology in a butterfly that has recently expanded its geographic range. *Oecologia*, 121(2):165–170, 1999.

[82] D.K. Holdsworth and A.F. Mark. Water and nutrient input: output budgets: effects of plant cover at seven sites in upland snow tussock grasslands of eastern and central otago, new zealand. *Journal of the Royal Society of New Zealand*, 20(1):1–24, 1990.

[83] Robert D. Holt. On the evolutionary ecology of species' ranges. *Evolutionary Ecology Research*, 5(2):159–178, 2003.

[84] Robert D. Holt and Michael Barfield. Trophic interactions and range limits: the diverse roles of predation. *Proceedings of the Royal Society of London B: Biological Sciences*, 276(1661):1435–1442, 2009.

[85] Thomas Hovestadt, Stefan Messner, and Hans Joachim Poethke. Evolution of reduced dispersal mortality and 'fat-tailed'dispersal kernels in autocorrelated landscapes. *Proceedings of the Royal Society of London B: Biological Sciences*, 268(1465):385–391, 2001.

[86] Clare L. Hughes, Jane K. Hill, and Calvin Dytham. Evolutionary trade-offs between reproduction and dispersal in populations at expanding range boundaries. *Proceedings of the Royal Society of London B: Biological Sci-*

ences, 270(Suppl 2):S147–S150, November 2003.

[87] Eric Imbert and Ophélie Ronce. Phenotypic plasticity for dispersal ability in the seed heteromorphic crepissancta (asteraceae). *Oikos*, 93(1):126–134, 2001.

[88] Anna Jakobsson and Ove Eriksson. A comparative study of seed number, seed size, seedling size and recruitment in grassland plants. *Oikos*, 88(3):494–502, 2000.

[89] Anna Jakobsson and Ove Eriksson. Trade-offs between dispersal and competitive ability: a comparative study of wind-dispersed asteraceae forbs. *Evolutionary Ecology*, 17(3):233–246, 2003.

[90] Jill E. Jankowski, Scott K. Robinson, and Douglas J. Levey. Squeezed at the top: interspecific aggression may constrain elevational ranges in tropical birds. *Ecology*, 91(7):1877–1884, 2010.

[91] B. Karlsson and A. Johansson. Seasonal polyphenism and developmental trade-offs between flight ability and egg laying in a pierid butterfly. *Proceedings of the Royal Society B: Biological Sciences*, 275(1647):2131–2136, Sep 2008.

[92] Victor Kasulo. 10. the impact of invasive species in african lakes. *The Economics of Biological Invasions*, page 183, 2000.

[93] Jon E. Keeley. Role of fire in seed germination of woody taxa in california chaparral. *Ecology*, 68(2):434–443, 1987.

[94] Elizabeth G. King and Derek A. Roff. Modeling the evolution of phenotypic plasticity in resource allocation in wing-dimorphic insects. *The American Naturalist*, 175(6):702–716, 2010.

[95] Mark Kirkpatrick and N. H. Barton. Evolution of a species' range. *The American Naturalist*, 150:1–23, Jul 1997.

[96] Eva Kisdi. Dispersal polymorphism in stable habitats. *Journal of Theoretical Biology*, 392:69–82, 2016.

[97] Walter D. Koenig, Frank A. Pitelka, William J. Carmen, Ronald L. Mumme, and Mark T. Stanback. The evolution of delayed dispersal in cooperative breeders. *The Quarterly Review of Biology*, 67(2):111–150, 1992.

[98] Hanna Kokko and Jan Ekman. Delayed dispersal as a route to breeding: territorial inheritance, safe havens, and ecological constraints. *The American Naturalist*, 160(4):468–484, 2002.

[99] Andrei N. Kolmogorov. Étude de l'équation de la diffusion avec croissance de la quantité de matière et son application à un problème biologique. *Bull. Univ. Moskow, Ser. Internat., Sec. A*, 1:1–25, 1937.

[100] Mark Kot, Mark A. Lewis, and P. van den Driessche. Dispersal data and the spread of invading organisms. *Ecology*, 77:2027–2042, Oct 1996.

[101] Alexander Kubisch, Tobias Degen, Thomas Hovestadt, and Hans Joachim Poethke. Predicting range shifts under global change: the balance between local adaptation and dispersal. *Ecography*, 36(8):873–882, 2013.

[102] Alexander Kubisch, Robert D. Holt, Hans-Joachim Poethke, and Emanuel A. Fronhofer. Where am i and why?: Synthesizing range biology and the eco-evolutionary dynamics of dispersal. *Oikos*, 123(1):5–22, Sep 2013.

[103] Patrick H. Leslie. Some further notes on the use of matrices in population mathematics. *Biometrika*, 35(3/4):213–245, 1948.

[104] P.H. Leslie and J.C. Gower. The properties of a stochastic model for the predator-prey type of interaction between two species. *Biometrika*, 47(3/4):219–234, 1960.

[105] Richard Levins. Some demographic and genetic consequences of environmental heterogeneity for biological control. *American Entomologist*, 15(3):237–240, 1969.

[106] Mark A. Lewis. Variability, patchiness, and jump dispersal in the spread of an invading population. *Spatial Ecology. The Role of Space in Population Dynamics and Interspecific Interactions*, pages 46–69, 1997.

[107] Mark A. Lewis and P Kareiva. Allee dynamics and the spread of invading organisms. *Theoretical Population Biology*, 43(2):141–158, 1993.

[108] Mark A. Lewis, Bingtuan Li, and Hans F. Weinberger. Spreading speed and linear determinacy for two-species competition models. *Journal of Mathematical Biology*, 45(3):219–233, Sep 2002.

[109] Guo Lin and Wan-Tong Li. Asymptotic spreading of competition diffusion systems: The role of interspecific competitions. *European Journal of Applied Mathematics*, 23(06):669–689, Aug 2012.

[110] S.D. Ling. Range expansion of a habitat-modifying species leads to loss of taxonomic diversity: a new and impoverished reef state. *Oecologia*, 156(4):883–894, 2008.

[111] Julie L. Lockwood, Martha F. Hoopes, and Michael P. Marchetti. *Invasion Ecology*. John Wiley & Sons, 2013.

[112] Sarah Lowe, Michael Browne, Souyad Boudjelas, and Maj De Poorter. *100 of the world's worst invasive alien species: a selection from the global inva-*

sive species database, volume 12. Invasive Species Specialist Group Auckland, 2000.

[113] Winsor H. Lowe, Gene E. Likens, and Bradley J. Cosentino. Self-organisation in streams: the relationship between movement behaviour and body condition in a headwater salamander. *Freshwater Biology*, 51(11):2052–2062, 2006.

[114] Alan F. Mark and JM Katharine Dickinson. Maximizing water yield with indigenous non-forest vegetation: a new zealand perspective. *Frontiers in Ecology and the Environment*, 6(1):25–34, Feb 2008.

[115] MATLAB. *version 8.3.0 (R2014a)*. The MathWorks Inc., Natick, Massachusetts, 2014.

[116] MATLAB. *version 9.0.0 (R2016a)*. The MathWorks Inc., Natick, Massachusetts, 2016.

[117] Angus R. Mcintosh, Barbara L. Peckarsky, and Brad W. Taylor. The influence of predatory fish on mayfly drift: extrapolating from experiments to nature. *Freshwater Biology*, 47(8):1497–1513, 2002.

[118] Mark A. McPeek and Robert D. Holt. The evolution of dispersal in spatially and temporally varying environments. *The American Naturalist*, 140(6):1010–1027, 1992.

[119] Alexander B. Medvinsky, Maria M. Gonik, Bai-Lian Li, and Horst Malchow. Beyond bt resistance of pests in the context of population dynamical complexity. *Ecological Complexity*, 4(4):201–211, 2007.

[120] Alexander B. Medvinsky, Andrew Y. Morozov, Vassili V. Velkov, Bai-Lian Li, Mikhail S. Sokolov, and Horst Malchow. Modeling the invasion of re-

cessive bt-resistant insects: an impact on transgenic plants. *Journal of Theoretical Biology*, 231(1):121–127, 2004.

[121] Gregor Mendel. *Experiments in plant hybridisation*. Harvard University Press, 1965.

[122] Diomar Cristina Mistro, Luiz Alberto Díaz Rodrigues, and Sergei Petrovskii. Spatiotemporal complexity of biological invasion in a space-and time-discrete predator–prey system with the strong allee effect. *Ecological Complexity*, 9:16–32, 2012.

[123] Simon Mole and Anthony J. Zera. Differential allocation of resources underlies the dispersal-reproduction trade-off in the wing-dimorphic cricket, gryllus rubens. *Oecologia*, 93(1):121–127, Feb 1993.

[124] Shiro Nakao and Chisako Chikamori. Temperature-dependent wing dimorphism in a japanese strain of tobacco thrips, frankliniella fusca (thysanoptera: Thripidae). *Applied Entomology and Zoology*, 48(3):337–343, 2013.

[125] Michael G. Neubert and Hal Caswell. Demography and dispersal: calculation and sensitivity analysis of invasion speed for structured populations. *Ecology*, 81(6):1613–1628, 2000.

[126] Kristjan Niitepõld, Anniina L.K. Mattila, Philip J. Harrison, and Ilkka Hanski. Flight metabolic rate has contrasting effects on dispersal in the two sexes of the glanville fritillary butterfly. *Oecologia*, 165(4):847–854, 2011.

[127] Kristjan Niitepõld, Alan D. Smith, Juliet L. Osborne, Don R. Reynolds, Norman L. Carreck, Andrew P. Martin, James H. Marden, Otso Ovaskainen, and Ilkka Hanski. Flight metabolic rate and pgi genotype

influence butterfly dispersal rate in the field. *Ecology*, 90(8):2223–2232, 2009.

[128] N.H. Ogden, A. Maarouf, I.K. Barker, M. Bigras-Poulin, L.R. Lindsay, M.G. Morshed, C.J. O'callaghan, F. Ramay, D. Waltner-Toews, and D.F. Charron. Climate change and the potential for range expansion of the lyme disease vector ixodes scapularis in canada. *International Journal for Parasitology*, 36(1):63–70, 2006.

[129] A. Okubo, P. K. Maini, M. H. Williamson, and J. D. Murray. On the spatial spread of the grey squirrel in britain. *Proceedings of the Royal Society B: Biological Sciences*, 238(1291):113–125, Nov 1989.

[130] Otso Ovaskainen and Stephen J. Cornell. Asymptotically exact analysis of stochastic metapopulation dynamics with explicit spatial structure. *Theoretical Population Biology*, 69(1):13–33, 2006.

[131] Rajendra K. Pachauri, Myles R. Allen, Vicente R. Barros, John Broome, Wolfgang Cramer, Renate Christ, John A. Church, Leon Clarke, Qin Dahe, Purnamita Dasgupta, et al. *Climate change 2014: synthesis report. Contribution of Working Groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. IPCC, 2014.

[132] Camille Parmesan and Gary Yohe. A globally coherent fingerprint of climate change impacts across natural systems. *Nature*, 421(6918):37–42, Jan 2003.

[133] Murray C. Peel, Brian L. Finlayson, and Thomas A. McMahon. Updated world map of the köppen-geiger climate classification. *Hydrology and Earth System Sciences Discussions*, 4(2):439–473, 2007.

[134] Sergei Petrovskii, Andrew Morozov, and Bai-Lian Li. On a possible origin of the fat-tailed dispersal in population dynamics. *Ecological Complexity*,

5:146–150, Jun 2008.

[135] Benjamin L. Phillips, Gregory P. Brown, and Richard Shine. Life-history evolution in range-shifting populations. *Ecology*, 91(6):1617–1627, Jun 2010.

[136] Benjamin L. Phillips, Gregory P. Brown, Justin M. J. Travis, and Richard Shine. Reid's paradox revisited: The evolution of dispersal kernels during range expansion. *The American Naturalist*, 172(S1):S34–S48, July 2008.

[137] Benjamin L. Phillips, Gregory P. Brown, Jonathan K. Webb, and Richard Shine. Invasion and the evolution of speed in toads. *Nature*, 439(7078):803–803, Feb 2006.

[138] F.X. Picó, N.J. Ouborg, and J.M. Van Groenendael. Fitness traits and dispersal ability in the herb tragopogon pratensis (asteraceae): decoupling the role of inbreeding depression and maternal effects. *Plant Biology*, 5(5):522–530, 2003.

[139] David Pimentel, Lori Lach, Rodolfo Zuniga, and Doug Morrison. Environmental and economic costs of nonindigenous species in the united states. *Bioscience*, 50(1):53–65, Jan 2000.

[140] David Pimentel, Rodolfo Zuniga, and Doug Morrison. Update on the environmental and economic costs associated with alien-invasive species in the united states. *Ecological Economics*, 52(3):273–288, Feb 2005.

[141] Hans J. Poethke, Thomas Hovestadt, and Oliver Mitesser. Local extinction and the evolution of dispersal rates: causes and correlations. *The American Naturalist*, 161(4):631–640, 2003.

[142] Hans Joachim Poethke and Thomas Hovestadt. Evolution of density–and

patch–size–dependent dispersal rates. *Proceedings of the Royal Society of London B: Biological Sciences*, 269(1491):637–645, 2002.

[143] Hans Joachim Poethke, Wolfgang W. Weisser, and Thomas Hovestadt. Predator-induced dispersal and the evolution of conditional dispersal in correlated environments. *The American Naturalist*, 175(5):577–586, 2010.

[144] Clement Reid. *The Origin of the British flora*. Dulau & Company, 1899.

[145] Marcel Rejmánek. Invasive plant species and invasible ecosystems. *Invasive Species and Biodiversity Management*, pages 79–102, 1999.

[146] David M. Richardson, Nicky Allsopp, Carla M. D'Antonio, Suzanne J. Milton, and Marcel Rejmánek. Plant invasions–the role of mutualisms. *Biological Reviews*, 75(1):65–93, 2000.

[147] Derek A. Roff. The genetic basis of wing dimorphism in the sand cricket, gryllus firmus and its relevance to the evolution of wing dimorphisms in insects. *Heredity*, 57(2):221, 1986.

[148] Derek A. Roff and Daphne J. Fairbairn. Wing dimorphisms and the evolution of migratory polymorphisms among the insecta. *American Zoologist*, 31(1):243–251, 1991.

[149] David J. Rogers and Sarah E. Randolph. The global spread of malaria in a future, warmer world. *Science*, 289(5485):1763–1766, 2000.

[150] Ophélie Ronce. How does it feel to be like a rolling stone? ten questions about dispersal evolution. *Annu. Rev. Ecol. Evol. Syst.*, 38:231–253, 2007.

[151] Ophélie Ronce and Isabelie Olivieri. Evolution of reproductive effort in a metapopulation with local extinctions and ecological succession. *The*

American Naturalist, 150(2):220–249, 1997.

[152] S.P. Rushton, P.W.W. Lurz, J. Gurnell, P. Nettleton, C. Bruemmer, M.D.F. Shirley, and A.W. Sainsbury. Disease threats posed by alien species: the role of a poxvirus in the decline of the native red squirrel in britain. *Epidemiology & Infection*, 134(3):521–533, 2006.

[153] Marjo Saastamoinen. Mobility and lifetime fecundity in new versus old populations of the glanville fritillary butterfly. *Oecologia*, 153(3):569–578, 2007.

[154] Marjo Saastamoinen. Heritability of dispersal rate and other life history traits in the glanville fritillary butterfly. *Heredity*, 100(1):39, 2008.

[155] Marjo Saastamoinen, Greta Bocedi, Julien Cote, Delphine Legrand, Frédéric Guillaume, Christopher W. Wheat, Emanuel A. Fronhofer, Cristina Garcia, Roslyn Henry, Arild Husby, Michel Baguette, Dries Bonte, Aurélie Coulon, Hanna Kokko, Erik Matthysen, Kristjan Niitepõld, Etsuko Nonaka, Virginie M. Stevens, Justin M. J. Travis, Kathleen Donohue, James M. Bullock, and Maria del Mar Delgado. Genetics of dispersal. *Biological Reviews*, 93(1):574–599, Aug 2017.

[156] Ann K. Sakai, Fred W. Allendorf, Jodie S. Holt, David M. Lodge, Jane Molofsky, Kimberly A. With, Syndallas Baughman, Robert J. Cabin, Joel E. Cohen, and Norman C. Ellstrand. The population biology of invasive species. *Annual Review of Ecology and Systematics*, pages 305–332, 2001.

[157] Osvaldo E. Sala, Stuart F. Chapin, Juan J. Armesto, Eric Berlow, Janine Bloomfield, Rodolfo Dirzo, Elisabeth Huber-Sanwald, Laura F. Huenneke, Robert B. Jackson, Ann Kinzig, et al. Global biodiversity scenarios for the year 2100. *Science*, 287(5459):1770–1774, 2000.

[158] W. E. Seabloom, S. W. Harpole, J. O. Reichman, and D. Tilman. Invasion, competitive dominance, and resource use by exotic and native california grassland species. *Proceedings of the National Academy of Sciences*, 100(23):13384–13389, Oct 2003.

[159] R. Shine, G. P. Brown, and B. L. Phillips. An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences*, 108:5708–5711, Mar 2011.

[160] Richard Shine. The ecological impact of invasive cane toads (bufo marinus) in australia. *The Quarterly Review of Biology*, 85(3):253–291, 2010.

[161] Richard Shine, Gregory P. Brown, and Benjamin L. Phillips. An evolutionary process that assembles phenotypes through space rather than through time. *Proceedings of the National Academy of Sciences of the United States of America*, 108(14):5708–5711, Apr 2011.

[162] Daniel Simberloff, Andrea María Relva, and Martín Nuñez. Gringos en el bosque: Introduced tree invasion in a native nothofagus/austrocedrus forest. *Biological Invasions*, 4(1):35–53, 2002.

[163] Adam D. Simmons and Chris D. Thomas. Changes in dispersal during species' range expansions. *The American Naturalist*, 164:378–395, Sep 2004.

[164] Adam D. Simmons and Chris D. Thomas. Changes in dispersal during species' range expansions. *The American Naturalist*, 164:378–395, Sep 2004.

[165] Frances S. Sivakoff, William F. Morris, Erik T. Aschehoug, Brian R. Hudgens, and Nick M. Haddad. Habitat restoration alters adult butterfly morphology and potential fecundity through effects on host plant quality. *Ecosphere*, 7(11), 2016.

[166] J. G. Skellam. Random dispersal in theoretical populations. *Biometrika*, 38(1):196–218, Jun 1951.

[167] Gabriele Sorci, Manuel Massot, and Jean Clobert. Maternal parasite load increases sprint speed and philopatry in female offspring of the common lizard. *The American Naturalist*, 144(1):153–164, 1994.

[168] Anne E. Sorensen. Somatic polymorphism and seed dispersal. *Nature*, 276(5684):174, 1978.

[169] David J. Spiegelhalter, Nicola G. Best, Bradley P. Carlin, and Angelika Van Der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4):583–639, 2002.

[170] J.A. Stamps. Habitat selection by dispersers: integrating proximate and ultimate approaches. *Dispersal*, pages 230–242, 2001.

[171] S.C. Stearns. Life-history tactics: a review of the ideas. *The Quarterly review of biology*, pages 3–47, 1976.

[172] Jolene T. Sutton, Shinichi Nakagawa, Bruce C. Robertson, and Ian G. Jamieson. Disentangling the roles of natural selection and genetic drift in shaping variation at mhc immunity genes. *Molecular Ecology*, 20(21):4408–4420, 2011.

[173] R Core Team et al. R: A language and environment for statistical computing. 2013.

[174] Joshua J. Tewksbury, Douglas J. Levey, Meribeth Huizinga, David C. Haak, and Anna Traveset. Costs and benefits of capsaicin-mediated control of gut retention in dispersers of wild chilies. *Ecology*, 89(1):107–117, 2008.

[175] C. D. Thomas, E. J. Bodsworth, R. J. Wilson, A. D. Simmons, Z. G. Davies, M. Musche, and L. Conradt. Ecological and evolutionary processes at expanding range margins. *Nature*, 411:577–581, May 2001.

[176] Fiona J. Thomson, Angela T. Moles, Tony D. Auld, and Richard T. Kingsford. Seed dispersal distance is more strongly correlated with plant height than with seed mass. *Journal of Ecology*, 99(6):1299–1307, 2011.

[177] Wilfried Thuiller. Patterns and uncertainties of speciestextquotesingle range shifts under climate change. *Global Change Biology*, 10(12):2020–2027, Dec 2004.

[178] D.M. Tompkins, A.R. White, and M. Boots. Ecological replacement of native red squirrels by invasive greys driven by disease. *Ecology Letters*, 6(3):189–196, 2003.

[179] J. M. J. Travis, R. W. Brooker, and C. Dytham. The interplay of positive and negative species interactions across an environmental gradient: insights from an individual-based simulation model. *Biology Letters*, 1(1):5–8, Mar 2005.

[180] J. M. J. Travis, T. Munkemuller, O. J. Burton, A. Best, C. Dytham, and K. Johst. Deleterious mutations can surf to high densities on the wave front of an expanding population. *Molecular Biology and Evolution*, 24(10):2334–2343, Jul 2007.

[181] Justin M. J. Travis and Calvin Dytham. Dispersal evolution during invasions. *Evolutionary Ecology Research*, 4(8):1119–1129, 2002.

[182] Justin M.J. Travis and Calvin Dytham. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, volume 265, chapter The evolution of dispersal in a metapopulation: a spatially explicit, individual-based model,

pages 17–23. 1998.

[183] Justin M.J. Travis, Karen Mustin, Kamil A. Bartoń, Tim G. Benton, Jean Clobert, Maria M. Delgado, Calvin Dytham, Thomas Hovestadt, Stephen C. F. Palmer, Hans Van Dyck, and Dries Bonte. Modelling dispersal: an eco-evolutionary framework incorporating emigration, movement, settlement behaviour and the multiple costs involved. *Methods in Ecology and Evolution*, pages 628–641, 2012.

[184] Lindsay A. Turnbull, Mark Rees, and Michael J. Crawley. Seed mass and the competition/colonization trade-off: a sowing experiment. *Journal of Ecology*, 87(5):899–912, 1999.

[185] Loïc Valéry, Hervé Fritz, Jean-Claude Lefeuvre, and Daniel Simberloff. In search of a real definition of the biological invasion phenomenon itself. *Biological Invasions*, 10(8):1345–1351, 2008.

[186] Wim van Saarloos. Front propagation into unstable states. *Physics Reports*, 386:29–222, Nov 2003.

[187] Brian W. Van Wilgen. Scientific challenges in the field of invasive alien plant management. 2004.

[188] Brian W. Van Wilgen, Richard M. Cowling, and Chris J. Burgers. Valuation of ecosystem services. *BioScience*, 46(3):184–189, 1996.

[189] Lawrence D. Venable and Joel S. Brown. The selective interactions of dispersal, dormancy, and seed size as adaptations for reducing risk in variable environments. *The American Naturalist*, 131(3):360–384, 1988.

[190] Wolfgang Viechtbauer. Conducting meta-analyses in r with the metafor package. *Journal of Statistical Software*, 36(3), 2010.

[191] Peter M. Vitousek. Biological invasions and ecosystem processes: Towards an integration of population biology and ecosystem studies. *Oikos*, 57(1):7–13, Feb 1990.

[192] Helene C. Weigang and Éva Kisdi. Evolution of dispersal under a fecundity-dispersal trade-off. *Journal of Theoretical Biology*, 371:145–153, 2015.

[193] Hans F. Weinberger, Mark A. Lewis, and Bingtuan Li. Analysis of linear determinacy for spread in cooperative models. *Journal of Mathematical Biology*, 45(3):183–218, Sep 2002.

[194] Hans F. Weinberger, Mark A. Lewis, and Bingtuan Li. Anomalous spreading speeds of cooperative recursion systems. *Journal of Mathematical Biology*, 55:207–222, Feb 2007.

[195] Wolfgang W. Weisser, Christian Braendle, and Nicole Minoretti. Predator-induced morphological shift in the pea aphid. *Proceedings of the Royal Society of London B: Biological Sciences*, 266(1424):1175–1181, 1999.

[196] Mark Westoby, Michelle Leishman, and Janice Lord. Comparative ecology of seed size and dispersal. *Proceedings of the Royal Society of London B: Biological Sciences*, 351(1345):1309–1318, 1996.

[197] Raj Whitlock. Relationships between adaptive and neutral genetic diversity and ecological structure and functioning: a meta-analysis. *Journal of Ecology*, 102:857–872, Jun 2014.

[198] Raj Whitlock, Gavin B. Stewart, Simon J. Goodman, Stuart B. Piertney, Roger K. Butlin, Andrew S. Pullin, and Terry Burke. A systematic review of phenotypic responses to between-population outbreeding. *Environmental Evidence*, 2(1), 2013.

[199] Bruce A. Wilcox and Duane J. Gubler. Disease ecology and the global emergence of zoonotic pathogens. *Environmental Health and Preventive Medicine*, 10(5):263–272, 2005.

[200] Mark Williamson. *Biological Invasions*, volume 15 of *Population and Community Biology Series*. Springer Science & Business Media, 1996.

[201] Andrew J. Young, Anne A. Carlson, and Tim Clutton-Brock. Trade-offs between extraterritorial prospecting and helping in a cooperative mammal. *Animal Behaviour*, 70(4):829–837, 2005.

[202] Anthony J. Zera and Robert F. Denno. Physiology and ecologyof dispersal polymorphism in insects. *Annual Review of Entomology*, 42:207–230, Jan 1997.

[203] Anthony J. Zera and Lawrence G. Harshman. The physiology of life history trade-offs in animals. *Annual Review of Ecology and Systematics*, 32:95–126, 2001.

[204] Anthony J. Zera and A. Larsen. The metabolic basis of life history variation: genetic and phenotypic differences in lipid reserves among life history morphs of the wing-polymorphic cricket, gryllus firmus. *Journal of Insect Physiology*, 47(10):1147–1160, 2001.

[205] Ying Zhang, Kongming Wu, Kris A.G. Wyckhuys, and George E. Heimpel. Trade-offs between flight and fecundity in the soybean aphid (hemiptera: Aphididae). *Journal of Economic Entomology*, 102(1):133–138, 2009.

[206] Lewis H. Ziska, Dana M. Blumenthal, Brett G. Runion, Raymond E. Hunt, and Hilda Diaz-Soltero. Invasive species and climate change: an agronomic perspective. *Climatic Change*, 105(1-2):13–42, 2011.